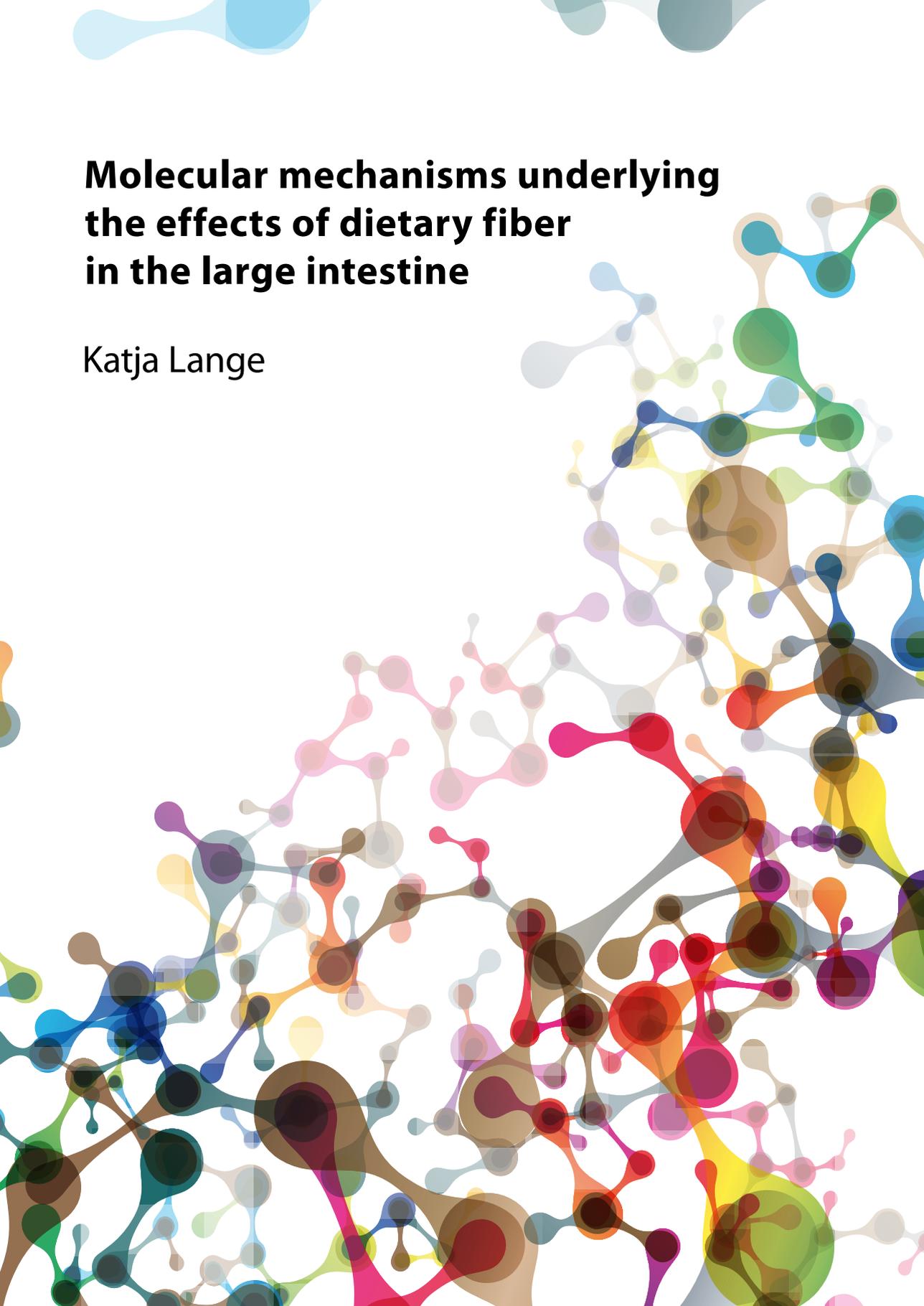


Molecular mechanisms underlying the effects of dietary fiber in the large intestine

Katja Lange



Thesis committee

Promotor

Prof. Dr Michael R. Müller
Professor of Nutrition, Metabolism and Genomics
Wageningen University

Co-promotor

Dr. Guido J.E.J. Hooiveld
Assistant professor, Division of Human Nutrition
Wageningen University

Other members

Prof. Dr Jerry Wells, Wageningen University
Dr Koen Venema, Maastricht University, Campus Venlo
Prof. Dr Albert K. Groen, Groningen University
Prof. John C. Mathers, Newcastle University, UK

This research was conducted under the auspices of the Graduate School VLAG (Advanced studies in Food Technology, Agrobiotechnology, Nutrition and Health Sciences).

Molecular mechanisms underlying the effects of dietary fiber in the large intestine

Katja Lange

Thesis

submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof Dr M.J. Kropff
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 24 April 2015
at 11 a.m. in the Aula.

Katja Lange

Molecular mechanisms underlying the effects of dietary fiber in the large intestine,
202 pages.

PhD thesis, Wageningen University, Wageningen, NL (2015)

With references, with summaries in Dutch and English

ISBN 978-94-6257-270-6

Abstract

Interactions between diet, microbiota and host response are important for intestinal health. Dietary fibers are known to promote intestinal health. Dietary fibers are edible plant-derived food components that encompass complex carbohydrates and lignin, resist the digestion in the small intestine of which some are degraded and fermented by gut microbiota in the large intestine, i.e. cecum and colon. The beneficial health effects of dietary fiber are suggested to be mediated by short-chain fatty acids (SCFA), which are produced by gut microbial fermentation. The underlying mechanisms of the interaction between dietary fiber, SCFA, and the host, however, are not in detail known.

The objective of the research described in this thesis was to investigate the molecular effects and mechanisms underlying the effects of dietary fiber and its fermentation products, SCFA, in the large intestine.

Firstly, the colonic transcriptional response to the main SCFA, acetate, propionate and butyrate, was investigated. SCFA were administered by rectal infusion in C57BL/6 mice fed a low fat/high carbohydrate (LFD) or high fat/low carbohydrate diet (HFD) and whole-genome gene expression analysis was performed on colonic scrapings by microarray technology. The analysis revealed specific and overlapping genes regulated between acetate, propionate and butyrate. In addition, gene response to SCFA was dependent on the diet, in particular for propionate. A set of propionate-regulated genes was activated on LFD while suppressed on a HFD and vice versa, indicating that diet composition is important factor in colonic response to SCFA.

Secondly, the molecular effects of different dietary fibers and a control diet on the large intestine were investigated. Five different dietary fibers (inulin, fructo-oligosaccharide, arabinoxylan, guar gum, resistant starch) and a control diet were fed to C57BL/6 mice (10 days). The transcriptional response to the fermentable fibers was comparable in gene expression, microbiota composition, and luminal SCFA level in colon. In common for all fermented dietary fibers, the transcriptional regulator Ppar γ was identified as potential upstream regulator for the mucosal gene expression response. Moreover, bacteria mainly belonging to *Clostridium* cluster XIVa were found to correlate with mucosal genes related to metabolic, energy-generating processes.

Next to common responses, analysis of the transcriptome revealed distinct responses of different dietary fibers. With respect to the cecal metatranscriptome, we identified distinct activities of bacterial families in the fermentation of dietary fiber. Moreover, using multivariate statistical analysis, we found correlations of the mucosal transcriptome with both the microbiota composition and metatranscriptome.

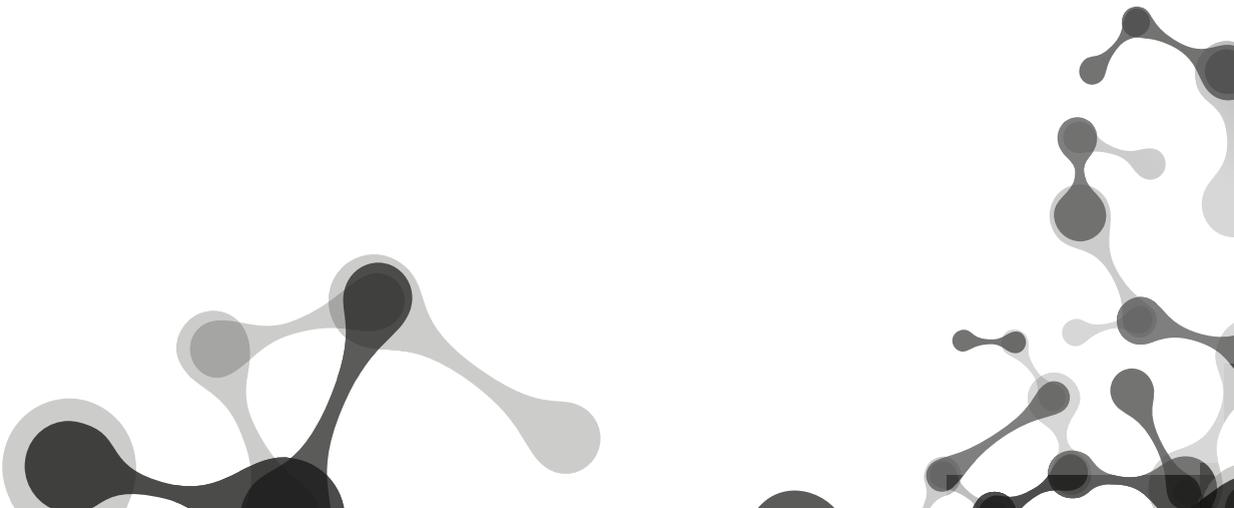
In addition, we showed that SCFA, particularly butyrate and to a lesser extend propionate, transactivate PPAR γ and regulate the PPAR γ target gene Angptl4 in colonic cells.

Thirdly, we tested the hypothesis that epithelial Ppar γ plays an important role in the fermentation of dietary fibers in the gut. Mice with an intestine-specific knock out (KO) of Ppar γ (cre-villin) and wild type (WT) mice were fed inulin (10 days). Whole-genome gene expression analysis of the colon revealed that diet had a larger effect than genotype on colonic, luminal microbiota composition, metabolome and mucosal transcriptome. We identified genes that were regulated by inulin in Ppar γ -dependent manner. In addition, we also identified genes regulated by butyrate in Ppar γ -dependent manner in organoids grown from colonic crypt cells derived from KO or WT mice.

In conclusion, we identified distinct mucosal gene expression responses to the main fermentation products of dietary fiber, SCFA, on both low fat/high carbohydrate and high fat/low carbohydrate diet backgrounds. Dietary fibers induce common and specific effects in colon. Epithelial Ppar γ partially governs the response to fermentation of dietary fiber in colon. Next to the commonalities of dietary fiber for intestinal physiology, specific and differential effects were identified for microbial gene activity and composition as well as mucosal transcriptome response indicating that omics tools are useful in elucidating and dissecting effects of dietary fiber.

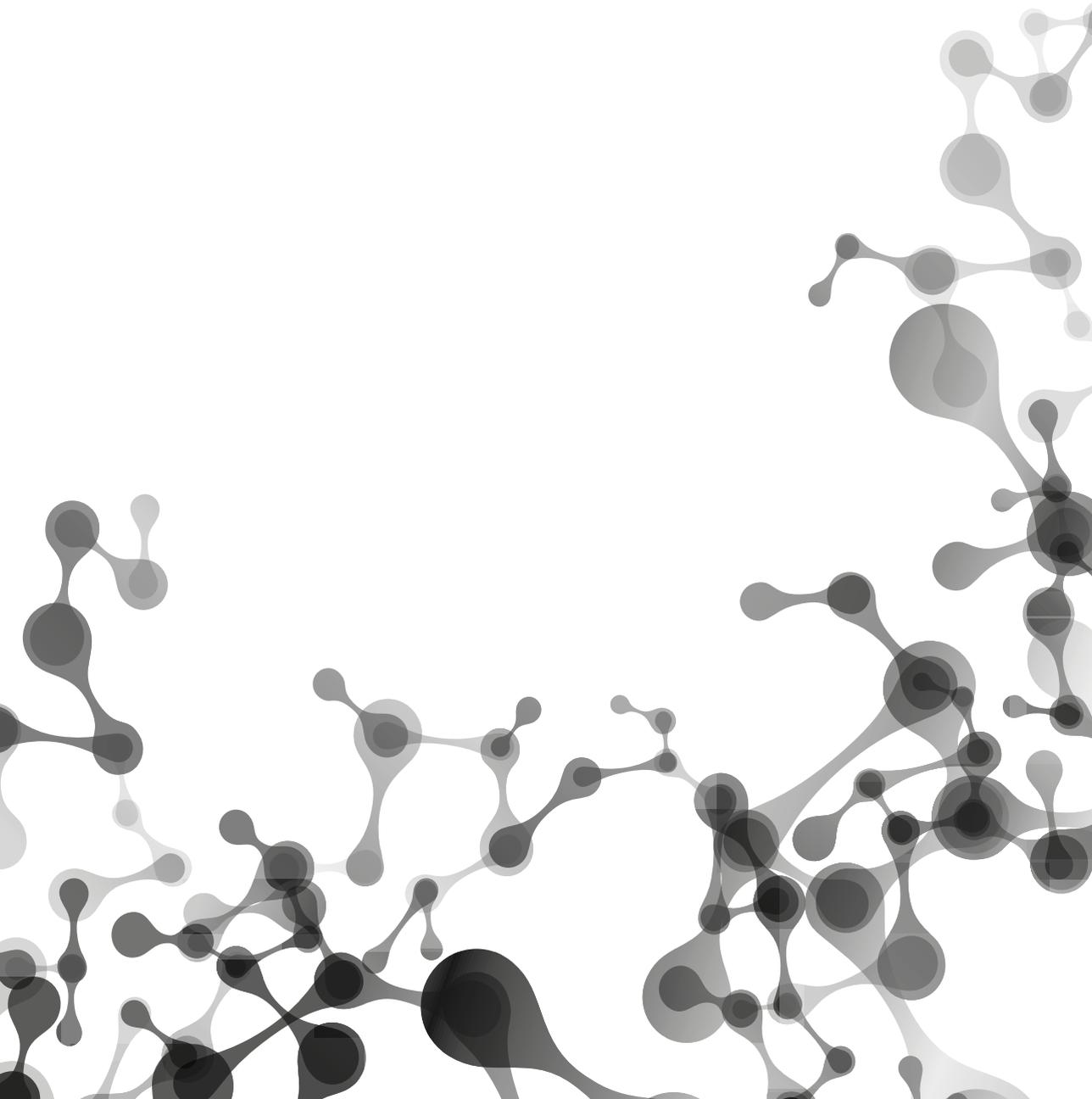
Contents

Chapter 1	General Introduction	9
Chapter 2	Transcriptomics revealed distinct short-chain fatty acid-induced changes on colonic gene expression in mice	23
Chapter 3	Comparison of the effects of five dietary fibers on mucosal transcriptional profiles and luminal microbiota composition and SCFA concentrations in murine colon	43
Chapter 4	Linking the fate of dietary fibers in the murine caecum to microbial transcriptome patterns	75
Chapter 5	Short chain fatty acids stimulate Angiopoietin-like 4 synthesis in human colon adenocarcinoma cells by activating PPAR γ	123
Chapter 6	Role of epithelial peroxisome proliferator-activated receptor (PPAR) γ in colonic response to fermentation of inulin in mice	153
Chapter 7	General Discussion	171
	Nederlandse samenvatting	183
	Acknowledgement	189
	About the author	195



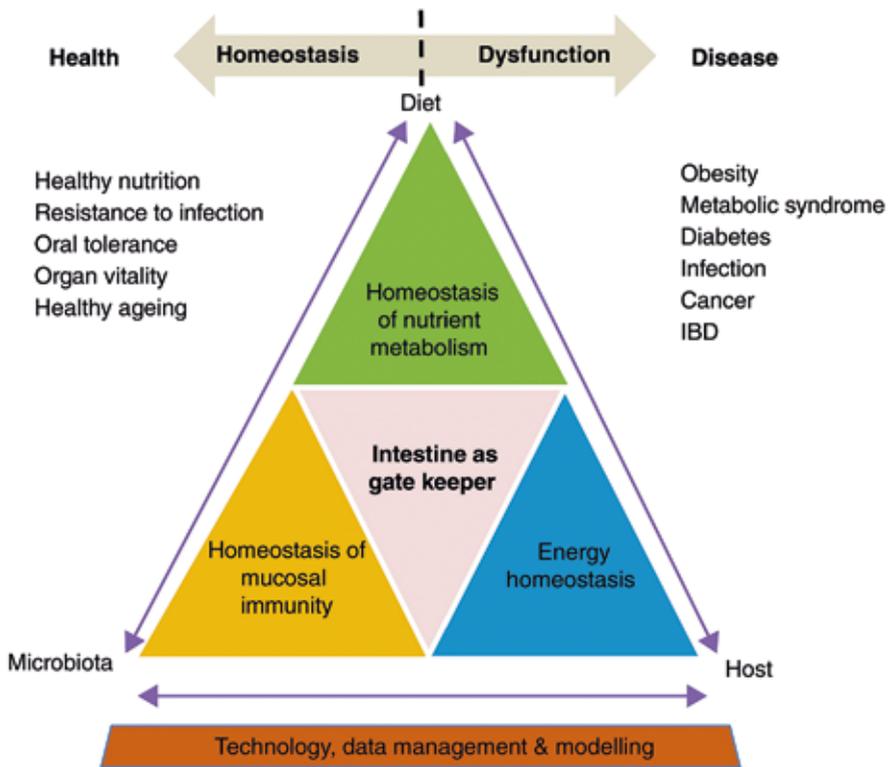
Chapter 1

General Introduction



The intestinal tract in mammals consists of small and large intestine. Its main function is the digestion and absorption of food and nutrients, which mainly takes place in the small intestine, whereas the large intestine is necessary for reabsorption of salts and water, elimination of waste products from the body, microbiota-associated synthesis of vitamins and fermentation thereby contributing to energy extraction from otherwise undigested food components such as dietary fiber.

“All disease begins in the gut” (Hippocrates). The gut can be regarded as gatekeeper of body homeostasis with the factors diet, microbiota and host (**Figure 1**). Beneficial interactions between these factors are essential for maintaining homeostasis. The underlying molecular mechanisms of this interaction, however, are largely unknown. One of the important dietary constituents required for keeping the balance are dietary fibers, that are in particular rich in plant food rich diets that are mainly fermented in the large intestine.



Current Opinion in Biotechnology

Figure 1 Model explaining the intestine as gatekeeper of body homeostasis [1]

Dietary fiber

In general, the term dietary fiber encompasses carbohydrates that are not digested in the small intestine and have health benefits. In more detail, these food components are quite diverse, in terms of physicochemical properties and in terms of health effects, hence the definition of dietary fiber has been rather challenging. Internationally recognized, the Codex Alimentarius Commission (CAC) adopted a definition from 2009:

Dietary fibre means carbohydrate polymers¹ with ten or more monomeric units², which are not hydrolysed by the endogenous enzymes in the small intestine of humans and belong to the following categories:

- *Edible carbohydrate polymers naturally occurring in the food as consumed,*
- *carbohydrate polymers, which have been obtained from food raw material by physical, enzymatic or chemical means and which have been shown to have a physiological effect of benefit to health as demonstrated by generally accepted scientific evidence to competent authorities,*
- *synthetic carbohydrate polymers which have been shown to have a physiological effect of benefit to health as demonstrated by generally accepted scientific evidence to competent authorities*

¹ *When derived from a plant origin, dietary fibre may include fractions of lignin and/or other compounds when associated with polysaccharides in the plant cell walls and if these compounds are quantified by the AOAC gravimetric analytical method (...). However, when extracted or even re-introduced into a food containing non digestible polysaccharides, they cannot be defined as dietary fibre.*

² *Decision on whether to include carbohydrates from 3 to 9 monomeric units should be left to national authorities*

Inclusion of dietary fibers in the daily diet is getting more and more attention because they are associated with the prevention of type 2 diabetes, obesity, colon cancer, cardiovascular health, body weight and appetite control, and promote intestinal health [2]. Dietary fibers are traditionally classified according to solubility in water. High soluble dietary fibers are generally linked to beneficial effects for prevention of gastrointestinal disorders and cardiovascular diseases [3], while insoluble fibers are generally attributed to laxative effects by increasing stool weight. The physiological effects linked to the term solubility are, however, not consistent and this classification might hence not be suitable for understanding physiological effects of dietary fibers [2,4]. Other physicochemical properties by which dietary fibers can be classified are water holding capacity, viscosity, binding of organic molecules and fermentability [5].

Fermentable dietary fibers are degraded by the intestinal microbiota and lead to the production of microbial metabolites such as short-chain fatty acids (SCFA). The health benefits of dietary fibers are generally attributed to the increased availability of SCFA. Different highly fermentable dietary fibers lead to different ratios of microbial metabolites

(SCFA) [6], which have in turn also different effects on the colon. Dietary fibers are complex food components, which have in addition to common also unique physicochemical properties and hence can be expected to have unique effects on the gut. Properties such as fermentability are dependent on the composition of the intestinal microbiota, which is an important player in the health effects mediated by diet to the host and hence cannot be ignored when studying dietary fibers.

Intestinal Microbiota

The human body is said to consist of approximately 10^{13} eukaryotic animal cells and is outnumbered by bacteria by a factor of 10 [7] of which the highest density is found in the gastrointestinal tract with up to 10^{12} /g of bacteria. The fact that certain bacteria influence health or cause disease has long been proposed and in recent years the importance of the complex ecosystem of intestinal bacteria as environmental factor for host (metabolic) health is becoming evident. First evidence came from studies with gnotobiotic mice, i.e. mice with a defined microbiota composition, including germ-free mice. Fed a chow diet (a diet high in carbohydrates), germ-free mice did not gain as much body fat as conventionalized mice [8], i.e. mice harboring microbiota and were protected from diet-induced obesity (with a high-fat, high-sugar diet) [9]. In addition, studies with transplantations of intestinal microbiota demonstrated the crucial role of microbiota for host health. For example, transplantation of microbiota from obese to non-obese gnotobiotic mice induced a phenotype in the recipient similar to that of the donor [10]. Similarly, transferring microbiota from twins discordant for obesity into germ-free mice resulted in a phenotype similar to that of the donor [11], which further supports the important role of microbiota for determining host metabolic phenotype. Moreover, in humans it was demonstrated that fecal transplantations by infusion of small intestinal microbiota from a lean donor into subjects with metabolic syndrome improved glucose handling in these subjects [12].

The composition of the intestinal microbiota thus plays an important role for health, because changes in microbiota community can influence its function for the host in harvesting energy, providing nutrients, shaping immunity and organ development [13]. A dysfunction between microbiota composition and host is associated with a variety of disorders not only restricted to the intestine (e.g. Crohn's disease [14]), but were also linked to metabolic syndrome [16] and psychological disorders (e.g. autism [15]). Hence, the intestinal microbiota is important for the gut as gatekeeper of whole body homeostasis [1]. It is well known that diet also plays an important role for whole body homeostasis and diet has a large impact on microbiota composition [17], [18], [19]. For example, microbiota composition was seen to be more different between mice fed a different diet than mice with different genotype status [20]. Animal-based diets have a high impact on composition shifts with higher abundance in bile acid-tolerant bacteria (e.g. *Bacteroides*) and decreased abundance of saccharolytic (carbohydrate fermenting) Firmicutes such as *Roseburia*, *Eubacterium rectale* and *Ruminococcus bromii* [18]. Saccharolytic bacteria can degrade

carbohydrates to produce microbial metabolites such as pyruvate and others, which can be used to form SCFA. Hence, SCFA potentially mediate the effect of microbiota and diet on whole body homeostasis.

Short-chain fatty acids (SCFA) - from microbiota to host

After degradation and fermentation of dietary fibers by intestinal microbiota, short-chain fatty acids (SCFA) are produced. SCFA, also known as volatile fatty acids, are organic acids of carbon chain length C1-6. The most abundant SCFA that are formed are acetate C2, propionate C3 and butyrate C4 (in total ca. 90% of SCFA), while valerate C5, hexanoate C6 and the branched-chain SCFA iso-butyrate C4 and iso-valerate C5 are present in lower amounts in colon [21]. SCFA are absorbed by the small intestine [22] and large intestinal epithelium [23]. Uptake across epithelial intestinal membrane takes place either via diffusion of the protonated SCFA or via active transport when dissociated. Next to bicarbonate exchange, monocarboxylate transport (via MCT1) and sodium-dependent monocarboxylate transport (via SMCT1) have been described as modes of transport [24]. All three SCFA are found in hepatic, portal and venous blood, with propionate being mainly taken up by the liver [25]. Propionate was found to affect cholesterol, fatty acid synthesis [26] and glucose metabolism [27], whereas acetate induces cholesterol synthesis [28]. Butyrate is mainly used by colonic epithelium where it serves as the preferred energy source as was demonstrated in isolated intestinal cells [29]. Many studies with butyrate have been performed in cultured cells demonstrating its anti-tumor effect through inhibition of cell proliferation and induction of apoptosis [30]. The molecular mechanisms underlying these effects of butyrate are widely attributed to its HDAC (histone deacetylase) inhibitory characteristics [31]. HDAC inhibition is related to regulation of gene expression by influencing chromatin structure and hence butyrate is important for gene regulation. Furthermore, SCFA are known to act via GPCR (G-protein coupled receptor) signaling. GPR41 (FFAR3) and GPR43 (FFAR2) have been identified as SCFA receptors [32,33]. The receptors play a role for regulation of immune response induced by SCFA [34], [35], [36]. Protective effects of SCFA on diet-induced obesity, however, seem to be independent of this receptor [37]. Hence, other, currently unknown mechanisms likely play a role.

Angiopoietin-like protein 4 (Angptl4)

Another proposed mechanism by which microbiota influences host energy metabolism is through induction of intestinal Fiaf/Angptl4 (fasting-induced adipose factor /angiopoietin-like 4) [8], [9]. Angptl4 belongs to angiopoietin-like family of proteins and has been identified as target of PPAR (peroxisome proliferator-activated receptor) [38]. Angptl4 is a circulating factor inhibiting lipoprotein lipase activity thereby affecting fat storage. The *Angptl4* mRNA levels are suppressed in small intestine in mice after conventionalization [8] and mice lacking Angptl4 were not protected from diet-induced obesity [9]. Hence, secreted, intestinal Angptl4 was suggested to play a role in protection from diet-induced

obesity by targeting distant tissues, which is, however, debated [39]. Fleissner et al. [39] showed that there was no difference in circulating *Angptl4* in GF mice on high-fat diet or western diet, despite higher *Angptl4* mRNA level compared to conventionalized mice. Detailed mechanisms by which *Angptl4* is regulated by microbiota are scarce.

Techniques to study microbiota

While knowledge about microbiota has traditionally been limited to culture-based studies, it is now increased with culture-independent methods that allow better characterization of gut microbiota by analysing bacterial genomes. These analyses can be achieved with different methods such as sequencing of 16S rRNA gene fragments and with a phylogenetic microarray. Originally developed for human species [40], phylogenetic microarrays have been developed and used for pig [41] and mouse as well [42]. The latter is referred to MITChip (mouse intestinal tract chip) as was used in this thesis. This microarray chip consists of oligonucleotide probes targeting two hypervariable regions of the 16S rRNA gene targeting 96 phylotypes. While the latter method can answer questions on microbiota composition ('Who is there'), metatranscriptomics using RNA sequencing can be used to answer questions on transcriptional activity of the microbiota ('What are they doing').

Nutrigenomics – studying host response

Studying interactions of nutrition and genes by using ~omics, high-throughput tools is collectively known as nutrigenomics. These tools include transcriptomics, proteomics and metabolomics. While transcriptomics measures RNA transcripts in a cell, proteomics assesses the protein products and metabolomics the metabolites produced. Nutrigenomics is successfully used to unravel underlying mechanisms of regulation and control of metabolic homeostasis by diet and host genetics [43]. Important in the control of metabolic homeostasis by diet is target gene regulation via transcriptional factors which act as nutrient sensors [44]. A well-studied group of nutrient sensors are PPARs. PPARs are ligand activated nuclear receptors of which the subtypes PPAR- α , - β , - γ have been identified. Next to synthetic ligands such as hypolipidemic drugs, also eicosanoids and fatty acids serve as ligands [45]. To study control and regulation of metabolic homeostasis by diet, nutrigenomics uses tools such as gene expression profiling and knock out mice models. Gene expression profiling can be performed with microarrays. Thereby, RNA isolated from tissue or cells is labeled with a dye for hybridization to an array that probes the expression of >22,000 genes. To test hypotheses generated from such gene expression profiling experiments, knockout mice models are useful. For example, using such nutrigenomics tools, the basic role of the isoform *Ppara* in small intestine [46] and in mediating the effect of dietary lipids on expression of genes related to barrier function has been demonstrated [47].

Omics data analysis

Omics data are large (high-throughput) amounts of data that are generated in nutrigenomics experiments. Guidelines on how to analyse these data, for example from microarray experiments, are given in [48]. In first instance, genes differentially regulated between treatment and control condition are identified by applying fold change and statistical significant cut offs. These differentially regulated genes can further be analysed to gain insight into potentially regulated biological processes. Different analysis tools are available to study over-represented biological processes in the selected set of genes, such as in Ingenuity, Enrichr, DAVID, ClueGO. While thereby sets of genes are chosen for the analyses (e.g. by significance and/or fold change cut off), other methods can be applied that include all genes in the analyses. The most-used method for such an analysis is gene-set enrichment analysis (GSEA) [49]. Thereby, a gene list is generated in which all analysed genes are ranked by e.g. fold change or t-test value, taking into account the variation within a treatment group. To explore the functional implications of the most regulated genes in this list, the enrichment of sets of genes with annotated biological functions is calculated. Information about the biological processes or other potential functional implications is derived from literature. In addition, to generate new hypotheses about potential functional implications and to better understand complex systems, multivariate statistical analyses can to be used. Multivariate statistical tests can explore relationships between different types of data, i.e. different omics data sets. MixOmics (integrOmics) is an R package available to perform such analysis like Partial Least Squares regression and correlation analyses and to deliver graphical outputs helping in the interpretation of relationships between two omics datasets derived from individual samples rather than groups of samples [50].

Outline of thesis

The work presented in this thesis was part of a larger collaborative project from the Netherlands Consortium for Systems Biology and the Top Institute for Food and Nutrition (NCSB/TIFN) in which partners from the Division of Human Nutrition, the Laboratory of Microbiology (Wageningen University) and Groningen University worked together addressing questions to SCFA formation by gut microbiota and subsequent metabolism in gut and liver in mice. The main part of this thesis focuses on the question to what extent and by which mechanisms do dietary fiber and its degradation products, SCFA, in the gut affect host transcriptional response. In addition, the results described in this thesis include results from the subproject addressing questions on microbial population dynamics in relation to exogenous factors (diet) and endogenous factors (host response).

First, the transcriptional response of colonic mucosal cells to different SCFA on the background of different diets was determined. Mice were fed either low or high fat diet

and treated by rectal infusion with a solution of one of the SCFA (acetate, propionate, butyrate) or a control (saline) for 6 consecutive days. The colonic mucosal cell scrapings were used for analysing whole-genome gene expression changes by microarray (**chapter 2**).

The objective of **chapter 3** and **4** was to investigate the effect of different dietary fibers on the gut. Therefore, mice were fed diets supplemented with one of the five dietary fibers, inulin, fructo-oligosaccharides, resistant starch, guar gum or arabinoxylan for 10 days. SCFA were determined by gas chromatography and whole-genome gene expression of colon was analysed by microarray. In addition, microbiota composition was measured with a phylogenetic microarray (MITChip) (**chapter 3**) and microbiota activity was determined in cecal content by sequencing to determine how different dietary fibers impact bacterial gene activity (**chapter 4**).

A potential factor mediating effects of diet and microbiota on host energy metabolism is ANGPTL4. Therefore, in **chapter 5** the mechanism of transcriptional regulation of colonic ANGPTL4 by SCFA was investigated. To investigate the mechanism, effects of SCFA on regulation of ANGPTL4 were tested in colonic cell lines, and with PPAR transactivation and coregulator binding assays. To further explore regulation of PPAR γ target genes by SCFA, an important regulator of adipogenesis, an adipocyte cell line was used. Furthermore, PPAR γ target gene regulation was explored by gene expression profiling of colonic cell lines treated with either butyrate or a specific, full PPAR γ agonist, Rosiglitazone, after which microarray was performed.

Results from **chapter 3** and **5** suggested that PPAR γ plays an important role in mediating the transcriptional response to dietary fiber and SCFA in colon. To test this hypothesis, intestine-specific *Ppary* knock out (KO) mice were fed with the dietary fiber inulin which was found to activate most target genes compared to other dietary fibers. Ppary-dependent gene expression changes were determined by microarray (**chapter 6**). In addition, to study the effect of butyrate on Ppary-dependent gene regulation, colonic organoids derived from *Ppary* KO and wild type mice were treated with butyrate and gene expression changes were measured by microarray (**chapter 6**).

Finally, in **chapter 7** the outcomes of **chapters 2** to **6** are discussed.

References

1. Martins dos Santos V, Muller M, de Vos WM (2010) Systems biology of the gut: the interplay of food, microbiota and host at the mucosal interface. *Curr Opin Biotechnol* 21: 539-550.
2. Slavin J (2013) Fiber and prebiotics: mechanisms and health benefits. *Nutrients* 5: 1417-1435.
3. Slavin JL, Savarino V, Paredes-Díaz A, Fotopoulos G (2009) A review of the role of soluble fiber in health with specific reference to wheat dextrin. *J Int Med Res* 37: 1-17.
4. Cummings JH, Stephen AM (2007) Carbohydrate terminology and classification. *Eur J Clin Nutr* 61 Suppl 1: S5-S18.
5. Blackwood AD, Salter J, Dettmar PW, Chaplin MF (2000) Dietary fibre, physicochemical properties and their relationship to health. *J R Soc Promot Health* 120: 242-247.
6. Henningsson AM, Bjorck IM, Nyman EM (2002) Combinations of indigestible carbohydrates affect short-chain fatty acid formation in the hindgut of rats. *J Nutr* 132: 3098-3104.
7. Savage DC (1977) Microbial ecology of the gastrointestinal tract. *Annu Rev Microbiol* 31: 107-133.
8. Backhed F, Ding H, Wang T, Hooper LV, Koh GY, et al. (2004) The gut microbiota as an environmental factor that regulates fat storage. *Proc Natl Acad Sci U S A* 101: 15718-15723.
9. Backhed F, Manchester JK, Semenkovich CF, Gordon JI (2007) Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. *Proc Natl Acad Sci U S A* 104: 979-984.
10. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, et al. (2006) An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 444: 1027-1031.
11. Ridaura VK, Faith JJ, Rey FE, Cheng J, Duncan AE, et al. (2013) Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science* 341: 1241-1244.
12. Vrieze A, Van Nood E, Holleman F, Salojarvi J, Kootte RS, et al. (2012) Transfer of intestinal microbiota from lean donors increases insulin sensitivity in individuals with metabolic syndrome. *Gastroenterology* 143: 913-916 e917.
13. Sommer F, Backhed F (2013) The gut microbiota--masters of host development and physiology. *Nat Rev Microbiol* 11: 227-238.
14. Gevers D, Kugathasan S, Denson LA, Vazquez-Baeza Y, Van Treuren W, et al. (2014) The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe* 15: 382-392.
15. Williams BL, Hornig M, Buie T, Bauman ML, Cho Paik M, et al. (2011) Impaired carbohydrate digestion and transport and mucosal dysbiosis in the intestines of children with autism and gastrointestinal disturbances. *PLoS One* 6: e24585.
16. Vijay-Kumar M, Aitken JD, Carvalho FA, Cullender TC, Mwangi S, et al. (2010) Metabolic syndrome and altered gut microbiota in mice lacking Toll-like receptor 5. *Science* 328: 228-231.
17. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen YY, et al. (2011) Linking long-term dietary patterns with gut microbial enterotypes. *Science* 334: 105-108.
18. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, et al. (2014) Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 505: 559-563.
19. Muegge BD, Kuczynski J, Knights D, Clemente JC, Gonzalez A, et al. (2011) Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* 332: 970-974.
20. Hildebrandt MA, Hoffmann C, Sherrill-Mix SA, Keilbaugh SA, Hamady M, et al. (2009) High-fat diet determines the composition of the murine gut microbiome independently of obesity. *Gastroenterology* 137: 1716-1724 e1711-1712.
21. Cook SJ, Sellin JH (1998) Review article: short chain fatty acids in health and disease. *Aliment Pharmacol Ther* 12: 499-507.
22. Kaji I, Iwanaga T, Watanabe M, Guth PH, Engel E, et al. (2014) SCFA transport in rat duodenum. *Am J Physiol Gastrointest Liver Physiol*: ajpgi 00298 02014.
23. McNeil NI, Cummings JH, James WP (1978) Short chain fatty acid absorption by the human large intestine. *Gut* 19: 819-822.
24. Miyauchi S, Gopal E, Fei YJ, Ganapathy V (2004) Functional identification of SLC5A8, a tumor suppressor down-regulated in colon cancer, as a Na(+)-coupled transporter for short-chain fatty acids. *J Biol Chem* 279: 13293-13296.

25. Cummings JH, Pomare EW, Branch WJ, Naylor CP, Macfarlane GT (1987) Short chain fatty acids in human large intestine, portal, hepatic and venous blood. *Gut* 28: 1221-1227.
26. Nishina PM, Freedland RA (1990) Effects of propionate on lipid biosynthesis in isolated rat hepatocytes. *J Nutr* 120: 668-673.
27. Wiltrout DW, Satter LD (1972) Contribution of propionate to glucose synthesis in the lactating and nonlactating cow. *J Dairy Sci* 55: 307-317.
28. Wong JM, de Souza R, Kendall CW, Emam A, Jenkins DJ (2006) Colonic health: fermentation and short chain fatty acids. *J Clin Gastroenterol* 40: 235-243.
29. Roediger WE (1982) Utilization of nutrients by isolated epithelial cells of the rat colon. *Gastroenterology* 83: 424-429.
30. Fung KY, Cosgrove L, Lockett T, Head R, Topping DL (2012) A review of the potential mechanisms for the lowering of colorectal oncogenesis by butyrate. *Br J Nutr* 108: 820-831.
31. Davie JR (2003) Inhibition of histone deacetylase activity by butyrate. *J Nutr* 133: 2485S-2493S.
32. Le Poul E, Loison C, Struyf S, Springael JY, Lannoy V, et al. (2003) Functional characterization of human receptors for short chain fatty acids and their role in polymorphonuclear cell activation. *Journal of Biological Chemistry* 278: 25481-25489.
33. Brown AJ, Goldsworthy SM, Barnes AA, Eilert MM, Tcheang L, et al. (2003) The Orphan G protein-coupled receptors GPR41 and GPR43 are activated by propionate and other short chain carboxylic acids. *J Biol Chem* 278: 11312-11319.
34. Kim MH, Kang SG, Park JH, Yanagisawa M, Kim CH (2013) Short-chain fatty acids activate GPR41 and GPR43 on intestinal epithelial cells to promote inflammatory responses in mice. *Gastroenterology* 145: 396-406 e391-310.
35. Maslowski KM, Vieira AT, Ng A, Kranich J, Sierro F, et al. (2009) Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. *Nature* 461: 1282-1286.
36. Trompette A, Gollwitzer ES, Yadava K, Sichelstiel AK, Sprenger N, et al. (2014) Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis. *Nat Med* 20: 159-166.
37. Lin HV, Frassetto A, Kowalik EJ, Jr., Nawrocki AR, Lu MM, et al. (2012) Butyrate and propionate protect against diet-induced obesity and regulate gut hormones via free fatty acid receptor 3-independent mechanisms. *PLoS One* 7: e35240.
38. Kersten S, Mandart S, Tan NS, Escher P, Metzger D, et al. (2000) Characterization of the fasting-induced adipose factor FIAF, a novel peroxisome proliferator-activated receptor target gene. *J Biol Chem* 275: 28488-28493.
39. Fleissner CK, Huebel N, Abd El-Bary MM, Loh G, Klaus S, et al. (2010) Absence of intestinal microbiota does not protect mice from diet-induced obesity. *Br J Nutr* 104: 919-929.
40. Rajilic-Stojanovic M, Heilig HG, Molenaar D, Kajander K, Surakka A, et al. (2009) Development and application of the human intestinal tract chip, a phylogenetic microarray: analysis of universally conserved phylotypes in the abundant microbiota of young and elderly adults. *Environ Microbiol* 11: 1736-1751.
41. Haenen D, Zhang J, Souza da Silva C, Bosch G, van der Meer IM, et al. (2013) A diet high in resistant starch modulates microbiota composition, SCFA concentrations, and gene expression in pig intestine. *J Nutr* 143: 274-283.
42. Geurts L, Lazarevic V, Derrien M, Everard A, Van Roye M, et al. (2011) Altered gut microbiota and endocannabinoid system tone in obese and diabetic leptin-resistant mice: impact on apelin regulation in adipose tissue. *Front Microbiol* 2: 149.
43. Afman L, Muller M (2006) Nutrigenomics: from molecular nutrition to prevention of disease. *J Am Diet Assoc* 106: 569-576.
44. Müller M, Kersten S (2003) Nutrigenomics: goals and strategies. *Nat Rev Genet* 4: 315-322.
45. Krey G, Braissant O, L'Horsset F, Kalkhoven E, Perroud M, et al. (1997) Fatty acids, eicosanoids, and hypolipidemic agents identified as ligands of peroxisome proliferator-activated receptors by coactivator-dependent receptor ligand assay. *Mol Endocrinol* 11: 779-791.
46. Bunger M, van den Bosch HM, van der Meijde J, Kersten S, Hooiveld GJ, et al. (2007) Genome-wide analysis of PPARalpha activation in murine small intestine. *Physiol Genomics* 30: 192-204.
47. de Vogel-van den Bosch HM, Bunger M, de Groot PJ, Bosch-Vermeulen H, Hooiveld GJ, et al. (2008) PPARalpha-mediated effects of dietary lipids on intestinal barrier gene expression. *BMC Genomics* 9: 231.

48. Allison DB, Cui X, Page GP, Sabripour M (2006) Microarray data analysis: from disarray to consolidation and consensus. *Nat Rev Genet* 7: 55-65.
49. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102: 15545-15550.
50. Le Cao KA, Gonzalez I, Dejean S (2009) integrOmics: an R package to unravel relationships between two omics datasets. *Bioinformatics* 25: 2855-2856.





Chapter 2

Transcriptomics revealed distinct short-chain fatty acid-induced changes on colonic gene expression in mice

Katja Lange*, Daniëlle Haenen*, Shohreh Keshtkar, Michael Müller and Guido J.E.J. Hooiveld

*These authors contributed equally to this work

Submitted.

Abstract

Acetate, propionate, and butyrate are the main short-chain fatty acids (SCFA) produced in the colon as a result of microbial fermentation, and are believed to be important mediators of the gut microbiota on host metabolism. The aim of this study was to investigate the impact of these three SCFA on the colonic transcriptome in mice, and whether responses were modulated by the composition of the diet. Mice were fed a semi-synthetic low fat/high carbohydrate (LFD), or high fat/low carbohydrate (HFD) diet for two weeks, after which they were subjected to repeated rectal infusion of either acetate, propionate, butyrate, or saline (control) solutions for 6 consecutive days. Colonic epithelial cell scrapings were subjected to gene expression profiling, and analyzed using a range of bioinformatics tools. SCFA induced distinct changes on colonic gene expression. Macronutrient composition of the diet impacted the colonic gene expression profiles more than SCFA treatment. In particular, for propionate opposite gene expression responses were observed on the LFD versus the HFD diets, which was not the case for acetate and to lesser extent for butyrate. Biological implications of differentially expressed genes reflected regulation of several metabolic and cell cycle related processes. In conclusion, SCFA induced distinct gene expression responses that are modulated by the diet composition.



Background

The gut microbiota is highly important for metabolic homeostasis and health, as it was found to play a crucial role in the regulation of energy homeostasis and fat storage [1,2]. The first evidence supporting this statement was derived from studies in germ-free mice, showing that conventionally raised mice had 42% more total body fat compared to germ-free mice, and conventionalization of germ-free mice resulted in a 57% increase in total body fat [3]. Furthermore, germ-free mice were protected from the development of obesity after consuming a high fat diet [4]. A possible factor explaining this effect are the microbial metabolites, among which the short-chain fatty acids (SCFAs) are produced at the highest levels. The physiological ratio and concentrations of the three main SCFA in mice vary dependent on the diet and the microbiota composition [5]. High concentrations of SCFA in mice were reported for colon with 100 mmol/L [6]. The three main SCFAs produced in the colon are acetate, propionate and butyrate, and each SCFA has distinct biological properties and can affect metabolism and health in a specific way. SCFA have been reputed to have regulatory role in fatty acid, glucose and cholesterol metabolism [5]. Butyrate is almost entirely utilized by colonocytes as their preferred energy substrate and therefore only a small proportion reaches the systemic circulation [7,8]. It has been observed that colonic administration (enema) of 100mM butyrate enhances the maintenance of colonic homeostasis in healthy humans, by regulating fatty acid metabolism, electron transport and oxidative stress pathways on gene expression level [9] Acetate and propionate are less efficiently used by the colon and can serve as anabolic substrates in tissues other than colon [10-12]. The molecular mechanisms underlying the effects of all three SCFA on colonic tissue are less well understood and have, to our knowledge, not been studied in the same model. Since measuring whole genome gene expression has the potential to obtain new insights in molecular mechanism underlying nutritional effects on metabolic homeostasis [13], it is a good tool for studying effects of SCFA on colonic tissue. High fat diets have been efficiently used to study effects of diet on metabolic disturbances in mice [14], [4]. Furthermore, dietary fibers have been linked to the prevention of metabolic disturbances by modulating food intake and body weight in both humans and rodent studies [15]. An increasing amount of evidence suggests that the health benefits observed with dietary fiber consumption are partly due to the action of these SCFAs [16]. The modulating effects of SCFA on high-fat diet induced gene expression, however, is less well studied.

In the current study we examined the effect of high fat diet intervention on changes in gene expression profile in colonic epithelial cells after colonic acetate, propionate, and butyrate administration. To have a closer look at the effects of each individual SCFA, we administered the SCFAs separately in this experiment. The experiment was performed in mice that were fed either a low fat or a high fat diet before and during SCFA treatment.

Methods

Ethics statement

The institutional and national guidelines for the care and use of animals were followed and the experiment was approved by the Local Committee for Care and Use of Laboratory Animals at Wageningen University.

Animals and diets

Male C57BL/6J mice were purchased from Charles River Laboratories (Maastricht, the Netherlands) at 8 weeks of age. Mice were housed individually in a light- and temperature-controlled animal facility, with light on from 23:00h to 11:00h. Mice had free access to water and food, and 24h food intake was monitored in every cage. Mice received standard laboratory chow (RMH-B, Arie Blok, Woerden, the Netherlands) during the first 2 weeks after arrival. Subsequently, mice were switched to semi-synthetic diets, either containing 10 en% of fat (low fat/high carbohydrate diet; LFD; $n = 24$) or 45 en% of fat (high fat/low carbohydrate diet; HFD; $n = 24$). Diets were based on Research Diets formulas D12450B/D12451, with adaptations regarding type of fat (palm oil instead of lard) and carbohydrates to mimic the fatty acid and carbohydrate composition of the average human diet in Western societies. Diets were prepared by Research Diet Services (Wijk bij Duurstede, The Netherlands). The complete composition of the diets is shown in **Table 1**.

SCFA infusions

After 2 weeks on either the LFD or HFD, the 6-day treatment period started. Mice were assigned to one of 4 treatment groups: Control, Acetate, Propionate or Butyrate. On 6 consecutive days, the mice were mildly sedated with a mixture of isoflurane (1.5%), nitrous oxide (70%) and oxygen (30%) 2h before the start of the dark phase, where after they received a rectal infusion of the test solutions. At time of infusion, mice were kept under sedation. Mice received an 80 μ L saline solution (Control; $n = 6$ per diet group), or an 80 μ L saline solution containing 100 mM sodium acetate (Acetate; $n = 6$), 100 mM sodium propionate (Propionate; $n = 6$) or 100 mM sodium butyrate (Butyrate; $n = 6$). All solutions had a pH of 6.5 and were isotonic. The solutions were administered by inserting a gel loading tip 3 cm into the rectum and slowly pushing the solution out of the tip.

Sample collection

Four hours after the rectal infusion on day 6, colonic scrapings were collected from 4 mice per group (32 in total). To reduce the inter-individual variation in physiological state at time of tissue collection, mice were provided a restricted amount of their habitual food (approximately 20% of their average daily intake of LFD or HFD) 2h before tissue removal. Mice were anaesthetized with isoflurane, where after the colon was excised and the length was measured. The adhering fat around the colon was carefully removed, and the

colon was cut open longitudinally. The intestinal content was removed and the tissue was rinsed with phosphate buffered saline. Subsequently, the epithelial lining of the colon was scraped. These scrapings were collected in 1.5 mL Eppendorf tubes, which were immediately snap frozen in liquid nitrogen and stored at -80°C for subsequent RNA isolation.

Table 1 Composition of the experimental diets

	LFD		HFD	
	g%	kcal%	g%	kcal%
Protein	19	20	24	20
Carbohydrate	67	70	41	35
Fat	4	10	24	45
Total		100		100
kcal/g	3.85		4.73	
Ingredient	g	kcal	g	kcal
Casein, lactic	200	800	200	800
L-Cystine	3	12	3	12
Corn starch	427.2	1,709	72.8	291
Maltodextrin	100	400	100	400
Sucrose	172.8	691	172.8	691
Cellulose, BW200	50	0	50	0
Soybean oil	25	225	25	225
Palm oil	20	180	177.5	1,598
Mineral mix S10026*	10	0	10	0
Dicalcium phosphate	13	0	13	0
Calcium carbonate	5.5	0	5.5	0
Potassium citrate, 1 H ₂ O	16.5	0	16.5	0
Vitamin mix V10001*	10	40	10	40
Choline bitartrate	2	0	2	0
Total	1,055	4,057	858.1	4,057

LFD, low fat diet; HFD, high fat diet. *Mineral mix S10026 contains the following (g/kg mineral mix): 41.9 magnesium oxide, 257.6 magnesium sulfate·7H₂O, 259 sodium chloride, 1.925 chromium KSO₄·12H₂O, 1.05 cupric carbonate, 0.035 potassium iodate, 21 ferric citrate, 12.25 manganous carbonate, 0.035 sodium selenite, 5.6 zinc carbonate, 0.20 sodium fluoride, 0.30 ammonium molybdate·4H₂O, 399.105 sucrose. *Vitamin mix V10001 contains the following (g/kg vitamin mix): 0.80 retinyl palmitate, 1.0 cholecalciferol, 10 all-rac-a-tocopheryl acetate, 0.08 menadione sodiumbisulfite, 2.0 biotin (1.0%), 1.0 cyanocobalamin (0.1%), 0.20 folic acid, 3.0 nicotinic acid, 1.6 calcium pantothenate, 0.70 pyridoxine-HCl, 0.60 riboflavin, 0.60 thiamin-HCl, and 978.42 sucrose.

RNA isolation and quality control

Total RNA was isolated from colon samples using TRIzol reagent (Invitrogen, Breda, the Netherlands) according to the manufacturer's instructions, followed by RNA Cleanup using the RNeasy Micro kit (Qiagen, Venlo, the Netherlands). Concentrations and purity of RNA samples were determined on a NanoDrop ND-1000 spectrophotometer (Isogen Life Science, De Meern, the Netherlands). RNA quality was verified on an Agilent 2100 Bioanalyzer (Agilent Technologies, Amstelveen, the Netherlands) using 6000 Nano Chips according to the manufacturer's instructions. RNA was judged as suitable for array hybridization only if samples exhibited intact bands corresponding to the 18S and 28S ribosomal RNA subunits, and displayed no chromosomal peaks or RNA degradation products (RNA Integrity Number > 8.0).

Microarray hybridization

Samples were subjected to genome-wide expression profiling using Affymetrix Mouse Gene 1.1 ST arrays (Affymetrix, Santa Clara, CA). Total RNA (100 ng) was used for whole transcript cDNA synthesis using the Ambion WT expression kit (Life Technologies, Nieuwekerk a/d IJssel, the Netherlands) and subsequently labeled using the Affymetrix GeneChip WT Terminal Labeling Kit (Affymetrix, Santa Clara, CA). Samples were hybridized, washed, stained, and scanned on an Affymetrix GeneTitan instrument. Detailed protocols for array handling can be found in the GeneChip WT Terminal Labeling and Hybridization User Manual (Affymetrix; P/N 702808, Rev. 4) and are also available on request.

Microarray analysis

Packages from the Bioconductor project [17], integrated in an online pipeline [18], were used to analyze the array data. Various advanced quality metrics, diagnostic plots, pseudoimages, and classification methods were used to determine the quality of the arrays prior statistical analysis [19,20]. From the 32 samples, one colon sample, derived from a mouse on the LFD in the *Acetate* group, was excluded from analyses because the array was of insufficient quality. The 828,268 probes on the Mouse Gene 1.1 ST array were redefined utilizing current genome information from the Entrez Gene database (custom CDF v16) [21]. Array data have been submitted to the Gene Expression Omnibus, a database repository for gene expression data hosted at the NCBI, under accession number GSE48856. Normalized gene expression estimates were obtained from the raw intensity values using the Robust Multi-array Average (RMA) preprocessing algorithm available in the library 'AffyPLM' using default settings [22]. The Mouse Gene 1.1 ST array probes the expression of 21,225 unique genes. Differentially expressed genes (probe sets) were identified using linear models, applying moderated t-statistics that implemented empirical Bayes regularization of standard errors (library 'limma'). The moderated t-test statistic has the same interpretation as an ordinary t-test statistic, except that the standard errors have been moderated across genes, i.e. shrunk to a common value, using a Bayesian model [23].



To adjust for both the degree of independence of variances relative to the degree of identity and the relationship between variance and signal intensity, the moderated t-statistic was extended by a Bayesian hierarchical model to define an intensity-based moderated T-statistic (IBMT) [24]. IBMT improves the efficiency of the Empirical Bayes moderated t-statistics and thereby achieves greater power while correctly estimating the true proportion of false positives. Probe sets that satisfied the criterion of $P < 0.01$ were considered to be significantly regulated. The dataset was filtered to only include probe sets that were active (i.e. expressed) in at least 3 samples using the universal expression code (UPC) approach (UPC score > 0.50) [25]. This resulted in the inclusion of 8,655 of the 21,187 probe sets. Partial Least Square Discriminant Analysis (PLS-DA) was used to discriminate the different treatments based on gene profiles. Analyses were performed in R using the library mixOmics. Similarity of the effects of each SCFA on the LFD and HFD was assessed using the Rank-rank Hypergeometric Overlap (RRHO) algorithm [26]. RRHO is a threshold-free method which visualizes the overlap between two ranked lists of differentially expressed genes. In brief, genes differentially expressed between SCFA infusion and the control treatment were ranked according to their IBMT t-statistic. The hypergeometric P value of overlap was calculated to determine the significance of overlapping genes between two datasets which is visualized in a matrix. On the x-axis of the matrix, genes were ranked by their degree of differential regulation (top up to bottom down) between SCFA infusion on the LFD compared to the control infusion on the LFD. On the y-axis genes were ranked based on their expression change with SCFA treatment on the HFD compared to control.

Analysis of functional implications

Changes in gene expression were related to functional changes using gene set enrichment analysis (GSEA) [27]. GSEA has the advantage that it is unbiased, because no gene selection step is used, and a score is computed based on all genes in a gene set. Briefly, genes were ranked based on the IBMT-statistic and subsequently analyzed for over- or underrepresentation in predefined gene sets. Gene sets were retrieved from the expert-curated KEGG, Biocarta, Reactome and WikiPathways pathway databases. Only gene sets consisting of more than 15 and fewer than 500 genes were taken into account. Statistical significance of GSEA results was determined using 1,000 permutations. The Enrichment Map plugin for Cytoscape was used for visualization and interpretation of the GSEA results [28,29].

Statistics

Results on food intake and body weight were expressed as means \pm SEM, regarding $P < 0.05$ as statistically significant.

Results

Food intake and body weight development

At the start of the 2-week intervention on the LFD and HFD, mice were stratified into 2 groups based on body weight, that averaged 22.7 g in both groups. Mice fed the HFD consumed the same amount of food as mice receiving the LFD (3.11 ± 0.08 versus 3.09 ± 0.04 g per day, respectively). However, since the HFD contained more energy per gram, the caloric intake was significantly higher in mice fed the HFD compared to mice fed the LFD (14.8 ± 0.36 versus 11.9 ± 0.14 kcal per day, respectively; $P < 0.0001$). During the treatment period, caloric intake of mice on the HFD remained significantly higher compared to mice fed the LFD (**Table 2**). SCFA treatment did not significantly affect caloric intake within both diet groups. After the 2-week intervention, body weight of mice fed the HFD was significantly higher compared to mice fed the LFD (26.6 ± 0.31 versus 23.9 ± 0.24 g respectively; $P < 0.0001$), which is in line with the observed difference in caloric intake between the 2 diet groups. During the SCFA treatment period, body weight change was similar with all SCFA treatments (data not shown).

Diet composition is more related to variation in gene expression than SCFA infusion

After the SCFA treatment period, gene expression in colonic scrapings was profiled on microarrays. PLS-DA revealed that the mice were clustering on first axis by the macronutrient composition of the diet (**Figure 1**). Mice on the left hand side of the plot received HFD whereas mice clustering on the right hand side received LFD. In addition,

Table 2 Daily food intake during the SCFA treatment period in mice fed the LFD or HFD

		Daily food intake (g)	Daily food intake (kcal)
LFD	Control	3.03 ± 0.076	11.67 ± 0.294
	Acetate	2.98 ± 0.118	11.48 ± 0.453
	Propionate	3.08 ± 0.114	11.84 ± 0.439
	Butyrate	2.89 ± 0.076	11.14 ± 0.293
	Average all treatments	3.00 ± 0.049	11.54 ± 0.188
HFD	Control	2.71 ± 0.095	12.80 ± 0.450
	Acetate	2.70 ± 0.097	12.77 ± 0.457
	Propionate	2.57 ± 0.084	12.17 ± 0.400
	Butyrate	2.48 ± 0.100	11.73 ± 0.474
	Average all treatments	$2.61 \pm 0.047^{**}$	$12.37 \pm 0.224^{**}$

Values are means \pm SEM, $n = 6$ per diet group, ** indicates a significant difference between average daily food intake in mice fed the LFD compared to mice fed HFD ($P < 0.01$). LFD, low fat diet; HFD, high fat diet.

mice receiving propionate on the HFD were deviating from all other mice, whereas on a LFD mice receiving propionate were overlapping with butyrate treated mice in their gene expression profile. These data indicate the importance of the macronutrient composition on the gene expression response to SCFA in colon.

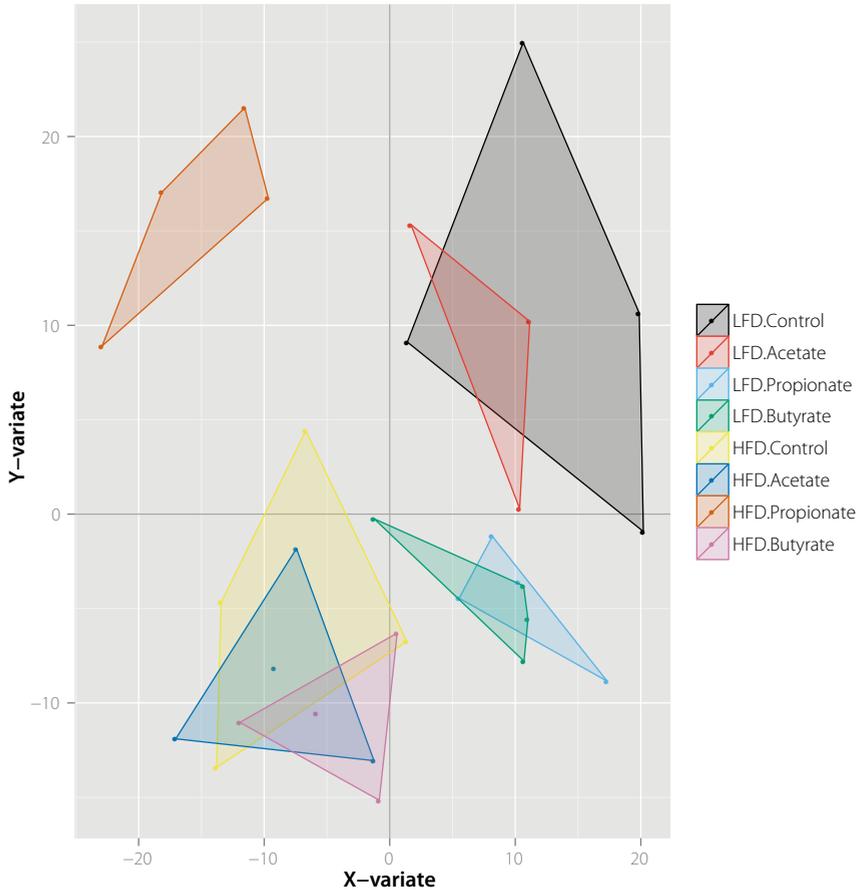


Figure 1 PLSDA plot colonic gene expression profile

Partial Least Square Discriminant Analysis (PLSDA) was used to discriminate samples of mice receiving SCFA or control treatment on the LFD or HFD based on their gene expression profile. The first two dimensions, explaining the highest variation, are shown.

SCFA-induced gene expression response is differentially modulated by diet

To characterize and compare the transcriptional responses to SCFA treatment on LFD and HFD, a cutoff free method, the RRHO analysis, was applied. As indicated by the RRHO p -value heatmaps (**Figure 2**), the transcriptional response to propionate was strikingly different on the two diets. For acetate we noticed in the RRHO heatmap a significant (yellow-red) area ranging from the lower left to the upper right corner of the map (**Figure 2A**). This showed that the gene ranking in both lists was rather similar, which indicated that the genes regulated by acetate on either the LFD or HFD were similar, even in case of a different extent of regulation. For butyrate treatment a comparable RRHO heatmap was

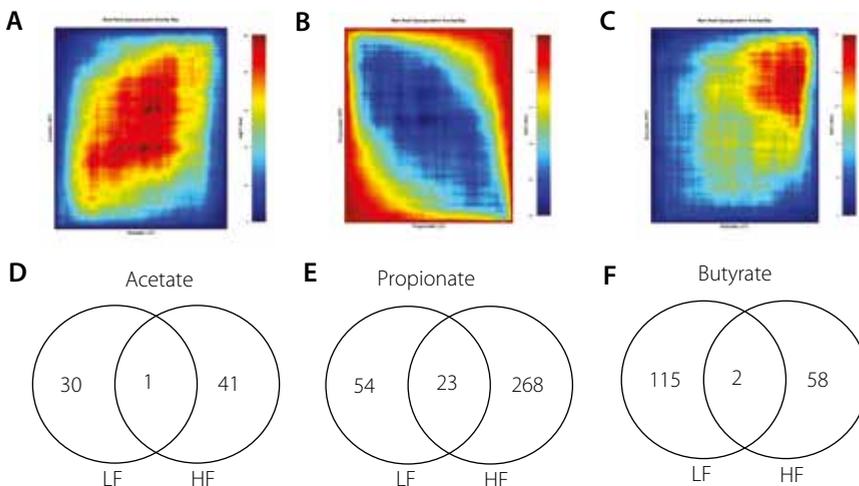


Figure 2 Effects of dietary fat on colonic gene expression

Heatmaps of Rank-Rank Hypergeometric Overlap (RRHO) analysis of the effect of 6-day acetate (**A**), propionate (**B**), and butyrate (**C**) infusion on colonic gene expression, and Venn diagrams showing the number of significantly regulated genes ($P < 0.01$) in the colon of mice after 6 days of rectal infusion of acetate (**D**), propionate (**E**) or butyrate (**F**). For RRHO analysis all 8655 filtered genes were used to determine the overlap between the transcriptional responses on the LFD and the HFD after SCFA infusion. The plots indicate the pattern of overlap between the LFD and the HFD background. Differentially regulated genes for each SCFA infusion on HFD versus LFD background were ranked based on Limma t value and plotted along the axis, with the X-axis representing the LFD and the Y-axis representing the HFD. Color represents the degree of hypergeometric distribution (Benjamini and Yekutieli-corrected signed \log_{10} transformed P value of overlap). Each Venn diagram shows the number of regulated genes on a LFD background, on a HFD background and the number of genes regulated on both dietary backgrounds. Effects of SCFA infusion were compared with effects after saline infusion in mice fed the same background diet (LFD or HFD).



observed, although the significance area was more pronounced in the top-right area (**Figure 2C**), which suggested a higher similarity between genes repressed by butyrate between the LFD and the HFD. In contrast, for propionate treatment a completely different RRHO heatmap was observed (**Figure 2B**). A significant (blue-green) area was observed mainly in the upper left quarter of the heatmap that further extended across to bottom right of the diagonal. Since RRHO conventions state a strong negative signal should be interpreted as a strong positive trend in the opposite sense [26], these results showed that genes induced by propionate on the LFD were suppressed on the HFD while genes induced by propionate on the HFD were suppressed on the LFD. The results indicate that diet composition has a major impact on propionate-induced gene regulation.

To further study effects of diet composition on SCFA-induced gene expression changes, we determined the number of genes regulated ($P < 0.01$) by acetate, propionate or butyrate treatment on LFD or HFD as compared to the control infusion (LFD or HFD, respectively). We observed an increase in the number of regulated genes for SCFA with increasing chain length under LFD conditions, i.e. acetate-LFD (31 genes), propionate-LFD (77 genes) and butyrate-LFD (117 genes) compared to control-LFD. This was not observed for HFD conditions. Propionate treatment resulted in a much higher number of regulated genes than acetate or butyrate on HFD (291 genes vs. 42 for acetate and 60 for butyrate, compared to HFD control). To compare genes regulated by each SCFA between the different background diets, we used Venn diagrams (**Figure 2D-F**). The comparison revealed that the number of genes commonly regulated by SCFA on the LFD or HFD, respectively was relatively low. For propionate we noticed that genes commonly regulated on LFD and HFD were regulated in opposite direction. From the 23 commonly regulated genes, 11 were induced on LFD and suppressed on HFD, and 11 genes were suppressed on the LFD and induced on HFD. Furthermore, we compared the number of genes regulated by the three SCFA per diet background (**Figure 3**) and revealed that SCFA induced very distinct gene expression responses with only one gene overlapping on both LFD or HFD, respectively.

In summary, the results showed that SCFA induced distinct responses in colonic gene expression and that diet can modulate the SCFA-induced gene expression response. The effect was most pronounced for propionate as was seen by the overlapping genes regulated in opposite direction under the two different diets.

Functional implications of gene expression changes

GSEA was performed to determine which biological processes were affected by the SCFA treatments. Results on the LFD and HFD were jointly visualized in enrichment maps (**Supplemental Figure 1A-C**). SCFA regulated various processes, and in particular for propionate these processes were majorly influenced by feeding LFD or HFD. Gene sets that were activated on LFD and suppressed on HFD with propionate infusion included DNA repair, cell cycle, metabolic (e.g. amino acid, gluconeogenesis) and vitamin related

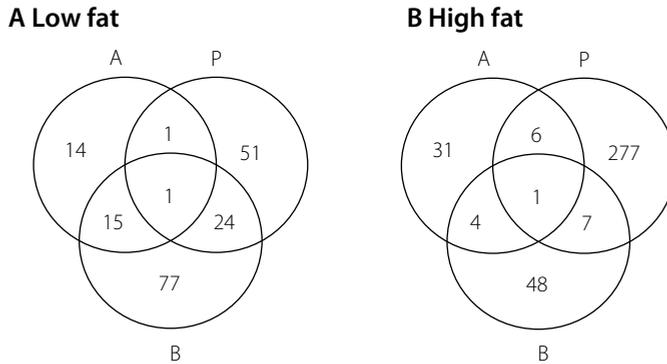


Figure 3 Venn diagrams of significantly regulated genes in colon

Each Venn diagram shows the number of significantly regulated genes ($P < 0.01$) in the colon of mice fed the LFD (**A**) or the HFD (**B**) after 6 days of rectal infusion of acetate, propionate, and butyrate. Effects of SCFA infusion were compared with effects after saline infusion in mice fed the same background diet (LFD or HFD).

processes, whereas gene sets suppressed on the LFD and activated on HFD encompassed translation and electron transport chain related processes. For acetate, no such oppositional regulation was observed. Acetate activated mainly gene sets related to translation, and suppressed mainly cell cycle, electron transport chain, mitochondrial translation and lipid digestion related processes. Similar to propionate, butyrate also regulated gene sets in opposite direction. Among the gene sets activated on HFD and suppressed on LFD by butyrate were mainly RNA, DNA and cell cycle related processes. In general, most gene sets regulated with butyrate on HFD were activated or not changed, while on LFD most gene sets were suppressed. Among these suppressed gene sets on LFD were PPAR signaling, lipoprotein metabolism and lipid digestion. On both diets, butyrate suppressed mainly electron transport chain and TCA cycle related gene sets.

Discussion

To our knowledge, this is the first experiment in which the impact of diets differing in macronutrient composition, in particular in fat content, on gene expression changes induced by colonic SCFAs was investigated. We showed that more variation in gene expression profiles was coinciding with macronutrient composition of the diet rather than with SCFA treatment. Gene expression responses to propionate were much dependent on the diet, whereas that was not observed for acetate and to lesser extend for butyrate. Functionally, expression changes reflected differential modulation of several metabolic processes.



Although a HFD combined with colonic SCFA treatment has not been studied before, experiments have been conducted to determine the effect of diets both high in fibers and fat. A mouse study reported that fermentable fiber decreased energy intake, body weight gain, and fat mass when added to a high fat diet [30]. A similar effect was observed in pigs fed diets differing in fat level (5% or 17.5%) and fiber type (fermentable or non-fermentable fiber) [31]. The fermentable fiber inulin resulted in an attenuation of the body weight development and fat mass accumulation induced by high fat feeding, suggesting a strong interplay between dietary fat and fiber type. We speculate this interaction might be, at least partially, attributed to an increased SCFA production resulting from fermentation of fibers by bacteria in the colon, as we showed that SCFA have modulatory effects on diet- induced gene expression.

It was previously shown that the colonic microbiota composition is affected by dietary fat content [32]. Although our experiment did not allow the measurement of microbiota composition, we assume the 2-week run-in period on either a low or high fat diet might have resulted in a stable change in microbiota composition, as has been reported [33-35]. Therefore, it is very likely that part of the colonic gene expression changes observed on the HFD background are the result of changed microbial composition. Interestingly, a number of studies showed that animals fed a HFD have lower intestinal concentrations of acetate, propionate, and butyrate than animals on a LFD [31,36]. It appears that HFD feeding reduces the fermentation capacity of the microbiota compared with a LFD, which might be related to the lower amount of carbohydrates in the HFD compared to LFD. In our study, the amount of colonic SCFAs that was infused was kept constant both in mice fed the LFD and in mice fed the HFD.

Effects of acetate and butyrate were comparable, whereas propionate induced very distinct changes in gene expression profile. For example, both acetate and butyrate suppressed gene sets related to lipid digestion, while this was not seen for propionate. Propionate specifically regulated gene sets related to amino acid metabolism. It is known that both acetate and butyrate enter the TCA cycle as acetyl-CoA, whereas propionate enters as succinyl-CoA [37]. It can be speculated that this pattern of SCFA utilization is related to the similar effects of acetate and butyrate on gene expression level. Furthermore, it is known that butyrate and acetate are interconverted by the intestinal microbiota [38]. This inter-conversion might have resulted in the production of additional acetate or butyrate in both infusion groups which might, in part, explain similar and overlapping effects on colonic gene expression. In general, the metabolic implications of lipoprotein and lipid metabolic processes are not well described for colonic tissue. However, the changes on gene expression level suggest a role for colonic tissue in lipid metabolism.

Conclusion

We showed that diet composition has a major impact on gene expression response to SCFAs in colon, an important metabolic organ, and conclude that effects of SCFA on gene

expression are modulated by diet-driven changes in colon. Next to butyrate, which has well known importance for colon, we found that also other SCFA have an impact on colonic gene regulation. While acetate and butyrate induced a similar gene expression profile, the effects of propionate on gene regulation are more specific.

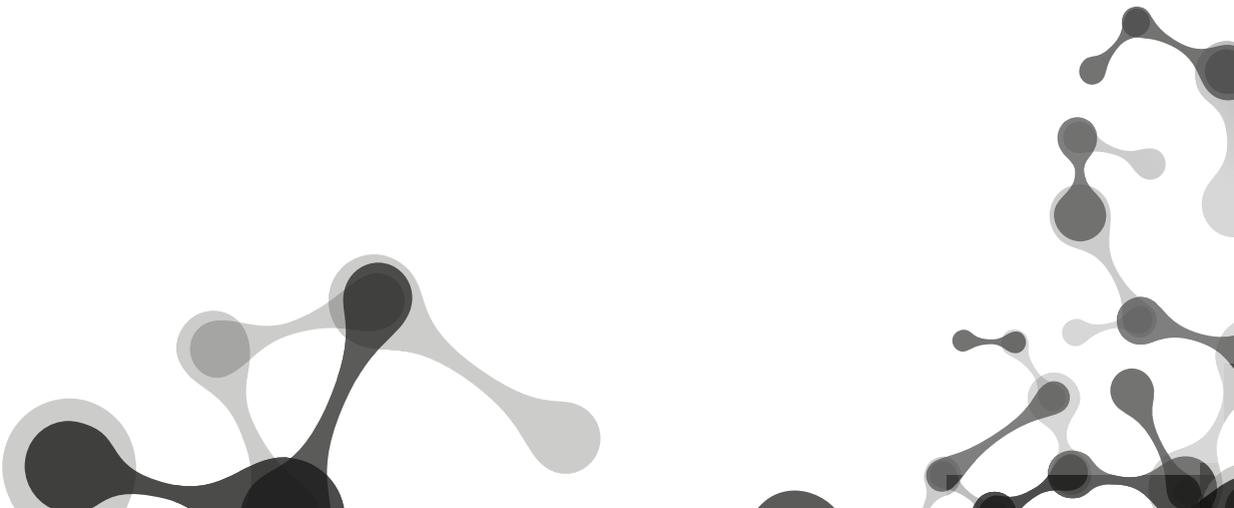
Acknowledgements

We would like to thank Jenny Jansen for performing the array hybridizations.

References

1. Grootaert C, Marzorati M, Van den Abbeele P, Van de Wiele T, Possemiers S (2011) Prebiotics to manage the microbial control of energy homeostasis. *Benef Microbes* 2: 305-318.
2. Delzenne NM, Neyrinck AM, Cani PD (2011) Modulation of the gut microbiota by nutrients with prebiotic properties: consequences for host health in the context of obesity and metabolic syndrome. *Microb Cell Fact* 10 Suppl 1: S10.
3. Backhed F, Ding H, Wang T, Hooper LV, Koh GY, et al. (2004) The gut microbiota as an environmental factor that regulates fat storage. *Proc Natl Acad Sci U S A* 101: 15718-15723.
4. Backhed F, Manchester JK, Semenkovich CF, Gordon JI (2007) Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. *Proc Natl Acad Sci U S A* 104: 979-984.
5. den Besten G, van Eunen K, Groen AK, Venema K, Reijngoud DJ, et al. (2013) The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J Lipid Res* 54: 2325-2340.
6. Alex S, Lange K, Amolo T, Grinstead JS, Haakonsson AK, et al. (2013) Short-chain fatty acids stimulate angiopoietin-like 4 synthesis in human colon adenocarcinoma cells by activating peroxisome proliferator-activated receptor gamma. *Mol Cell Biol* 33: 1303-1316.
7. Roediger WE (1982) Utilization of nutrients by isolated epithelial cells of the rat colon. *Gastroenterology* 83: 424-429.
8. Fleming SE, Fitch MD, DeVries S, Liu ML, Kight C (1991) Nutrient utilization by cells isolated from rat jejunum, cecum and colon. *J Nutr* 121: 869-878.
9. Vanhoutvin SA, Troost FJ, Hamer HM, Lindsey PJ, Koek GH, et al. (2009) Butyrate-induced transcriptional changes in human colonic mucosa. *PLoS One* 4: e6759.
10. Wong JM, de Souza R, Kendall CW, Emam A, Jenkins DJ (2006) Colonic health: fermentation and short chain fatty acids. *J Clin Gastroenterol* 40: 235-243.
11. Roy CC, Kien CL, Bouthillier L, Levy E (2006) Short-chain fatty acids: ready for prime time? *Nutr Clin Pract* 21: 351-366.
12. Bloemen JG, Venema K, van de Poll MC, Olde Damink SW, Buurman WA, et al. (2009) Short chain fatty acids exchange across the gut and liver in humans measured at surgery. *Clin Nutr* 28: 657-661.
13. Müller M, Kersten S (2003) Nutrigenomics: goals and strategies. *Nat Rev Genet* 4: 315-322.
14. Winzell MS, Ahren B (2004) The high-fat diet-fed mouse: a model for studying mechanisms and treatment of impaired glucose tolerance and type 2 diabetes. *Diabetes* 53 Suppl 3: S215-219.
15. Lattimer JM, Haub MD (2010) Effects of dietary fiber and its components on metabolic health. *Nutrients* 2: 1266-1289.
16. Topping DL, Clifton PM (2001) Short-chain fatty acids and human colonic function: roles of resistant starch and nonstarch polysaccharides. *Physiol Rev* 81: 1031-1064.
17. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5: R80.
18. Lin K, Kools H, de Groot PJ, Gavai AK, Basnet RK, et al. (2011) MADMAX - Management and analysis database for multiple ~omics experiments. *J Integr Bioinform* 8: 160.
19. Heber S, Sick B (2006) Quality assessment of Affymetrix GeneChip data. *OMICS* 10: 358-368.
20. Brettschneider J, Collin F, Bolstad BM, Speed TP (2008) Quality Assessment for Short Oligonucleotide Microarray Data. *Technometrics* 50: 241-264.
21. Dai M, Wang P, Boyd AD, Kostov G, Athey B, et al. (2005) Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res* 33: e175.
22. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, et al. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4: 249-264.
23. Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3: Article3.
24. Sartor MA, Tomlinson CR, Wesselkamper SC, Sivaganesan S, Leikauf GD, et al. (2006) Intensity-based hierarchical Bayes method improves testing for differentially expressed genes in microarray experiments. *BMC Bioinformatics* 7: 538.

25. Piccolo SR, Withers MR, Francis OE, Bild AH, Johnson WE (2013) Multiplatform single-sample estimates of transcriptional activation. *Proc Natl Acad Sci U S A* 110: 17778-17783.
26. Plaisier SB, Taschereau R, Wong JA, Graeber TG (2010) Rank-rank hypergeometric overlap: identification of statistically significant overlap between gene-expression signatures. *Nucleic Acids Res* 38: e169.
27. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102: 15545-15550.
28. Merico D, Isserlin R, Stueker O, Emili A, Bader GD (2010) Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One* 5: e13984.
29. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T (2011) Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27: 431-432.
30. Delmee E, Cani PD, Gual G, Knauf C, Burcelin R, et al. (2006) Relation between colonic proglucagon expression and metabolic response to oligofructose in high fat diet-fed mice. *Life Sci* 79: 1007-1013.
31. Yan H, Potu R, Lu H, Vezzoni de Almeida V, Stewart T, et al. (2013) Dietary fat content and fiber type modulate hind gut microbial community and metabolic markers in the pig. *PLoS One* 8: e59581.
32. Hildebrandt MA, Hoffmann C, Sherrill-Mix SA, Keilbaugh SA, Hamady M, et al. (2009) High-fat diet determines the composition of the murine gut microbiome independently of obesity. *Gastroenterology* 137: 1716-1724 e1711-1712.
33. Cani PD, Amar J, Iglesias MA, Poggi M, Knauf C, et al. (2007) Metabolic endotoxemia initiates obesity and insulin resistance. *Diabetes* 56: 1761-1772.
34. Cani PD, Neyrinck AM, Fava F, Knauf C, Burcelin RG, et al. (2007) Selective increases of bifidobacteria in gut microflora improve high-fat-diet-induced diabetes in mice through a mechanism associated with endotoxaemia. *Diabetologia* 50: 2374-2383.
35. Murphy EF, Cotter PD, Healy S, Marques TM, O'Sullivan O, et al. (2010) Composition and energy harvesting capacity of the gut microbiota: relationship to diet, obesity and time in mouse models. *Gut*.
36. Kimura I, Ozawa K, Inoue D, Imamura T, Kimura K, et al. (2013) The gut microbiota suppresses insulin-mediated fat accumulation via the short-chain fatty acid receptor GPR43. *Nat Commun* 4: 1829.
37. Black AL, Kleiber M (1958) The transfer of carbon from propionate to amino acids in intact cows. *J Biol Chem* 232: 203-209.
38. den Besten G, Lange K, Havinga R, van Dijk TH, Gerding A, et al. (2013) Gut-derived short-chain fatty acids are vividly assimilated into host carbohydrates and lipids. *Am J Physiol Gastrointest Liver Physiol*.



Chapter 3

Comparison of the effects of five dietary fibers on mucosal transcriptional profiles and luminal microbiota composition and SCFA concentrations in murine colon

Katja Lange*, Floor Hugenholtz*, Melliana C Jonathan,
Henk A Schols, Michiel Kleerebezem, Hauke Smidt, Michael Müller,
Guido J.E.J. Hooiveld

*These authors contributed equally to this work

Submitted.

Abstract

Consumption of diets rich in fibers has been associated with beneficial effects on gastrointestinal health. However, detailed, comparative studies on molecular effects of dietary fibers in colon are limited. The aim of our study was to investigate and compare the effects of five fibers on the mucosal transcriptome, together with alterations in the luminal microbiota composition and SCFA concentrations in colon. Mice were fed fibers that differed in carbohydrate composition or a control diet for 10 days. Colonic gene expression profiles and luminal microbiota composition were determined by microarray techniques, and integrated using multivariate statistics. Our data showed a distinct reaction of the host and microbiota to resistant starch, a fiber that was not completely fermented in the colon, whereas the other fibers induced similar responses on gene expression and microbiota. Consistent associations were revealed between fiber-induced enrichment of *Clostridium* cluster XIVa representatives and *Parabacteroides distasonis*, and changes in mucosal expression of genes related to energy metabolism. The nuclear receptor Ppar γ was predicted to be an important regulator of the mucosal responses. Taken together, our results suggest that despite different source and composition, fermentable fibers induce a highly similar mucosal response that may at least be partially governed by Ppar γ .

Introduction

From a scientific and regulatory perspective multiple definitions of dietary fiber have been proposed, but in 2009 the first internationally agreed definition was adopted by the CODEX Alimentarius Commission [1,2]. Essentially, dietary fibers encompass predominantly complex carbohydrate polymers with ten or more monomeric units that escape digestion and absorption in the small intestine. However, when derived from a plant origin, dietary fiber may include fractions of lignin and/or other compounds associated with polysaccharides in the plant cell walls, and the decision on whether to include carbohydrates of 3 to 9 monomeric units is ultimately left up to national authorities [1,2]. Importantly, consumption of fiber-rich diets has been associated with a variety of beneficial health effects, including the improvement of gastrointestinal homeostasis [3-6].

Upon entering the large intestine, dietary fibers are completely or partially fermented by the gut microbiota, resulting mostly in the generation of short chain fatty acids (SCFA) [7,8]. The main SCFA that are produced are acetate, propionate, and butyrate, but the SCFA production and ratio depend on the type of fiber, which potentially affect the existing microbiota composition. [9-11].

Much research has been conducted to investigate effects of dietary fibers on gut health and microbiota composition, however, most *in vitro*, animal or human studies investigated a single fiber at a time. Regarding the microbiota, both of the two main phyla Bacteroidetes and Firmicutes contain a large range of carbohydrate utilizing enzymes [12], indicating a role in complex dietary fiber degradation for both phyla. For example, it has been reported that *Bacteroides* spp. have a range of glycoside hydrolases and are capable of switching between different substrates [12,13], producing several metabolites including succinate, acetate and propionate [9]. On the other hand, some groups of bacteria, such as some members of the *Bifidobacterium* genus, are specialized to ferment certain oligosaccharides [14,15], with lactate and acetate as the main fermentation products [9]. As a result dietary fibers modulate the microbiota composition by triggering bacteria that directly feed on them, but also the cross-feeding bacteria that depend on these primary degraders. Data on the genome-wide transcriptional effects of dietary fibers in colonic mucosa are scarce. This is remarkable since there is a major interest in characterizing the genes and networks that are regulated by food components, because this contributes to our understanding of a healthy diet [16,17]. It has only been reported that differential gene expression due to consumption of resistant starch (RS) suggested improvement of structure and function of the gastrointestinal tract in rats [18], and induced catabolic but suppressed immune and cell division pathways in the proximal colon of pigs [19]. In addition, it was shown that oligofructose (FOS) induced expression of genes involved in the tricarboxylic acid (TCA) cycle, oxidative phosphorylation and proteasome-mediated degradation of intracellular proteins in the rat colon [20].

The aim of the current study was to comprehensively investigate and compare the effects of five different fibers on the mucosal transcriptome, together with alterations in the luminal microbiota composition in the murine colon. To this end, mice were fed diets enriched with fibers that differed in carbohydrate composition or a control diet for 10 days. Colonic gene expression profiles and luminal microbiota composition were determined by microarray techniques, and integrated using multivariate statistics.

Methods

Ethics statement

The institutional and national guidelines for the care and use of animals were followed and the experiment was approved by the Local Committee for Care and Use of Laboratory Animals at Wageningen University (DRS code: 2010167).

Animals, diets, design and sampling

Male C57BL/6J mice were purchased from Charles River Laboratories (Maastricht, the Netherlands) at 6 weeks of age. Mice were housed in pairs in a light- and temperature-controlled animal facility of Wageningen University (12 hour light-dark cycle; light on from 11h PM to 11h AM, 21 °C). Mice had free access to water and food throughout the entire experimental period. Upon arrival, mice were fed standard lab chow (RMH-B, Arie Blok, Woerden, the Netherlands) for 3 wks. Subsequently, all mice were adjusted to the control diet, a standard semi-synthetic low fat diet containing corn starch, for 2 wks. To achieve similar weight distribution among the diet groups, mice were stratified according to their body weight to one of the six diet groups (n=10 per diet group), i.e. control (CON), inulin (IN), oligofructose (FOS), arabinoxylan (AX), guar gum (GG) or resistant starch (RS). The diets enriched in fiber were identical to the control diet, except that 10% (w/w) of corn starch was replaced by each fiber (20% for RS, see below). Inulin (brand name Frutafit IQ), a fructan isolated from roots of chicory (DP2-60, average chain length monomers: 11) and oligofructose (brand name Frutalose OFF), another fructan (DP2-7) obtained from inulin by partial enzymatic hydrolysis, were a gift of Dr Diederick Meyer (Sensus, Roosendaal, the Netherlands). Arabinoxylan (brand name NAXUS), mechanically extracted from wheat endosperm, was a gift of Dr Hans van der Saag (BioActor, Maastricht, the Netherlands). Guar gum (brand name Viscogum), a galactomannan, and resistant starch (brand name C* ActiStar 11700), a retrograded starch obtained from tapioca by an enzymatic treatment, were obtained from Cargill R&D Centre Europe (Vilvoorde, Belgium). On the basis of physical and chemical characteristics, the RS used in this study can be classified as RS type 3 (RS3). According to the supplier, the RS was only 50% resistant to digestion in the small intestine, and was therefore included in the diet at a 20% (w/w) level, i.e. double the amount of the other fibers. Diets were prepared by Research Diet Services (Wijk bij Duurstede, The Netherlands). Detailed composition of the diets is presented in **Supplemental Table 1**.

Mice were fed the fiber or control diets for 10 days. This period was based on the observation in mice that a switch from a standard low-fat, plant polysaccharide-rich diet to a high-fat, high-sugar Western diet resulted in a shift in microbiota composition that stabilized by 7 days [21]. On the day of sections, mice were fasted for 4hrs (starting at 5 AM). Mice then received a calibrated meal of 1g of their habitual diet to reduce the inter-individual variation in physiological state at time of tissue collection. Four hours later mice were anaesthetized with isoflurane, and the colon was excised. The adhering fat around the colon was carefully removed, and the colon was cut open longitudinally. The luminal content was sampled and the tissue was rinsed with ice-cold phosphate buffered saline. Subsequently, the epithelial lining of the colon was scraped off. Luminal content and scrapings were collected in tubes, which were immediately snap frozen in liquid nitrogen and stored at -80°C.

RNA isolation, Affymetrix microarray processing and analysis

Colonic scrapings (n=6 per diet group) were subjected to genome-wide expression profiling. In brief, total RNA was isolated from epithelial scrapings and were hybridized on Mouse Gene 1.1 ST arrays (Affymetrix). Packages from the Bioconductor project [22], integrated in an online pipeline [23], were used for quality control and statistical analysis of the array data. Due to insufficient quality, 1 array from the control group had to be excluded from further analysis. The dataset was filtered to only include probe sets that were active (i.e. expressed) in at least 5 samples using the universal expression code (UPC) approach (UPC score > 0.50) [24]. This resulted in the inclusion of 8,831 (42%) of the 21,187 probe sets. Differentially expressed probe sets were identified by using linear models and an intensity-based moderated t-statistic [25,26]. Probe sets that satisfied the criterion of $P < 0.01$ were considered to be significantly regulated. Array data have been submitted to the Gene Expression Omnibus under accession number GSE59494. Detailed information on microarray processing and data analysis can be found under **Supplemental Methods**.

Biological interpretation of array data

Changes in gene expression were related to biologically meaningful changes using gene set enrichment analysis (GSEA) [27]. It is well accepted that GSEA has multiple advantages over analyses performed on the level of individual genes [27-29]. GSEA evaluates gene expression on the level of gene sets that are based on prior biological knowledge, e.g. published information about biochemical pathways or signal transduction routes, allowing more reproducible and more interpretable analysis of gene expression data; GSEA is unbiased, because no gene selection step (fold change and/or p-value cutoff) is used; a GSEA score is computed based on all genes in gene set, which boosts the signal-to-noise ratio and allows to detect affected biological processes that are due to only subtle changes in expression of individual genes. Gene sets were retrieved from the expert-curated KEGG, Biocarta, Reactome and WikiPathways pathway databases. Only

gene sets consisting of more than 15 and fewer than 500 genes were taken into account. For each comparison, genes were ranked on their t-value that was calculated by the empirical Bayes method. Statistical significance of GSEA results was determined using 1,000 permutations. The Enrichment Map plugin for Cytoscape was used for visualization and interpretation of the GSEA results [30]. Upstream Regulator Analysis in IPA (content version 18030641 released 2013; Ingenuity Systems) was used to identify the cascade of potential upstream transcriptional regulators that may explain the observed gene expression changes in the data set, and whether they are likely activated or inhibited. Biological interpretation of selected genes, e.g. identified by multivariate correlation analysis, was performed in Enrichr [31].

DNA isolation, microbiota MITchip processing and analysis

Total bacterial DNA was extracted from colonic luminal content samples (n=4-5 per diet group) using the repeated bead beating plus column method [32]. Quantification of the overall bacterial community density was performed by qPCR targeting the 16 rRNA gene, whereas the microbial community composition was analyzed using the Mouse Intestinal Tract Chip (MITChip) [33] (for further details also see **Supplemental Methods**). The relative abundance of 96 genus-level bacterial groups detected on the MITChip was determined by the Robust Probabilistic Averaging algorithm [34]. To assess the correlation of the microbial groups with all diets groups, multivariate redundancy analysis (RDA) was performed as implemented in Canoco for Windows 4.5 [35]. The Monte Carlo Permutation test was used to assess the significance of the variation in the dataset in relation to the diet.

Short-chain fatty acid analysis of colonic luminal content

Luminal samples (n=3-5 per diet group) were analyzed for SCFA concentration by gas chromatography as described before [36].

Saccharide analysis of colonic luminal content

Luminal samples (n=1 per group) were used for analysis of mono-, di- and oligosaccharides by high-performance anion-exchange chromatography coupled with pulsed amperometric detection (HPAEC-PAD), essentially as described before [37]. Full experimental details can be found under Supplemental Methods.

Multivariate integration and correlation analysis

To get insight into the interactions between changes in gene expression and microbiota composition, the datasets were combined using the linear multivariate method partial least squares (PLS) [38]. This analysis ignores diet group membership. For 15 mice both gene expression and microbiota composition data was available, but to increase power the dataset was expanded with 7 measurements performed in mice housed in the same cage. Since we did not want to make any 'a priori' assumption on the relationship between

the two sets of variables that were analyzed, the canonical correlation framework of PLS was used [39]. Both datasets were log₂ transformed before analysis, and the correlation matrices were visualized in clustered image maps [40]. Analyses were performed in R using the library mixOmics [41].

Results

Modulation of gene expression in colonic epithelial cells by dietary fibers

Expression profiling by microarray was performed to assess the genome-wide differences in gene expression in colonic epithelial cells of animals fed different dietary fibers. First individual mice were compared on basis of their gene expression profile and diet using sparse PLS-Discriminant Analysis (DA). Results of this analysis showed that the gene expression profiles were clearly distinguishable between samples of the control, RS and the other 4 fiber groups (FOS, IN, GG, AX) (**Figure 1**). Within the cluster of IN, FOS, AX, and GG samples, separation was much less distinct due to individual animal variation, but the AX samples tended to separate from the majority of IN, FOS and GG samples. Thus, based on their effect on global gene expression patterns, diet groups could be grouped in three main clusters. Next, the number of significantly ($P < 0.01$) regulated genes per fiber compared to control were determined. The expression of in total 1,733 genes was altered by at least 1 fiber. The largest number of changed genes was observed for FOS (925), and smallest for RS (287). The percentage of uniquely regulated genes varied between 21% (IN) and 40% (RS). Only 28 genes were commonly regulated by all fibers (**Supplemental Figure 1**). Taken together, RS appeared to be most similar to the control diet and induced only few, but specific gene expression changes, whereas the other fibers induced a larger, but more similar responses.

Biological interpretation of differential gene expression in epithelial cells

It is well recognized that analyses of expression data on the level of gene sets is more sensitive, reproducible and interpretable than the analysis at the level of individual genes (see Methods section). Gene set enrichment analysis (GSEA) was therefore performed to gain better insight into the biological functions represented by the regulated genes. Significantly induced or suppressed gene sets ($p < 0.001$, FDR < 0.01) were identified for each fiber (**Supplemental Figure 2**), and biological processes summarized in **Table 1**. For RS the smallest number of changed gene sets were found, and these sets were mostly distinct from the sets found for the other fibers, being in line with the expression data at the level of individual genes. Gene sets induced by IN, GG, FOS and AX were largely comparable. Among the induced gene sets were energy-generating processes such as TCA cycle, electron transport chain, and oxidative phosphorylation. In addition, glycolysis,

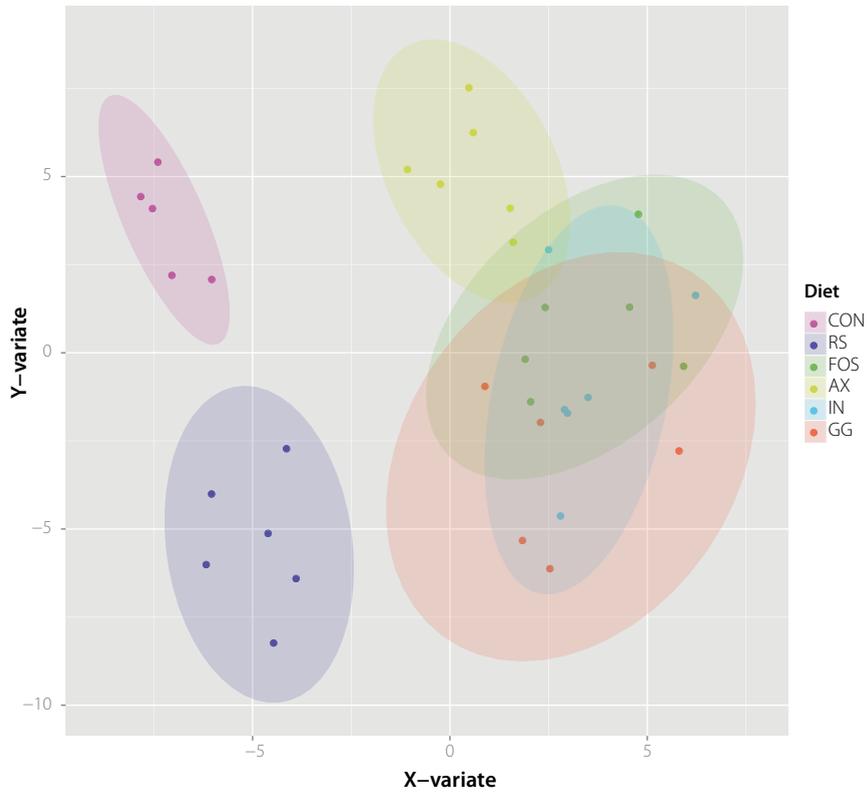


Figure 1 Effects of dietary fiber on gene expression profiles in colonic epithelial cells

PLS-DA score plot of gene expression profiles in colonic epithelial cells of mice fed 5 different fibers or the control diet. In the plot the samples (individual mice) were plotted based on the two main variates. Ellipses indicate 90% confidence intervals of the scores.

phase I/II metabolism and target genes of the transcription factors peroxisome proliferator-activated receptor (PPAR) and nuclear factor (erythroid-derived 2)-like 2 (NRF2), were commonly activated for the fiber diets, except RS. Suppression of gene sets describing RNA processing and translation was found for RS as well as IN, GG and AX. Fiber-specific responses included activation of fatty acid oxidation and adipocytokine signaling that was specifically found for IN, cholesterol/steroid synthesis and arachidonic and linoleic fatty acid metabolism, specific for GG, tryptophan metabolism (AX only), and branched-chain amino acid catabolism, degradation/proteasome system and mitochondrial translation (FOS only).

Table 1 Biological processes regulated by dietary fibers

Gene sets significantly regulated ($P < 0.001$, $FDR < 0.01$) by dietary fiber compared to control were determined by gene set enrichment analysis. Subsequently, clusters were manually grouped and labeled to highlight the prevalent biologic functions among related gene sets. See *Supplemental Figure 2* for a high-resolution version of the maps that includes the names of the gene sets.

	RS	FOS	AX	IN	GG
Transcription				↓	
Translation	↓		↓	↓	↓
RNA processing	↓		↓	↓	↓
Chromatin modification				↓	↓
DNA repair	↓			↓	
Glucose metabolism		↑	↑	↑	↑
Phase I/II metabolism	↓	↑	↑	↑	↑
PPAR & Nrf2 targets		↑	↑	↑	↑
TCA cycle & Oxidative phosphorylation		↑	↑	↑	↑
Lipid metabolism		↑		↑	↑
Amino acid metabolism		↑	↑		↑
		Degradation/ proteasome system	Tryptophan metabolism	Fatty acid oxidation	Cholesterol/ Steroid metabolism
		Branched chain amino acid catabolism/ Arginine, Proline metabolism		Adipocytokine signaling	Arachidonic & Linoleic acid metabolism
		Mitochondrial translation			
		UPR (unfolded protein response)			

Identification of Ppar γ as important upstream regulator

The underlying mechanisms by which the fibers modulated gene expression changes are not well understood. We therefore aimed to identify potential upstream transcriptional regulators that could explain the observed shifts in gene expression profiles (**Table 2**). In line with results obtained by GSEA, PPAR, particularly the isoform Ppar γ , was predicted to

Table 2 Common and specific potential upstream regulator in colon of mice after feeding different fiber diets as determined by Ingenuity Systems Pathway Analysis Software

(Transcriptional regulator and ligand-dependent nuclear receptor which showed an activation z-score ≥ 2 or ≤ -2 and a p-value < 0.05 are displayed)

	Activation score per dietary fiber				
	RS	FOS	AX	IN	GG
Ppary		2.83	2.01	4.23	3.07
HNF4A				2.58	3.50
TP53				2.36	2.82
ATF4		2.61			2.43
PPARGC1A				2.39	2.08
XBP1					2.93
NR5A2		2.61			
SREBF1					2.58
FOXC2				2.43	
SREBF2					2.22
PTTG1				2.21	
NR1H2					2.09
CEBPB				2.02	
KDM5B	2.00				
NCOA2				2.00	
TP63			-2.15		
STAT5B					-2.16
MBD2		-2.23			
STAT5A					-2.36
MYC				-2.63	

be activated by FOS, AX, GG, but most for IN. This was supported by the observation that many Ppary targets genes were indeed found to be regulated by one or more of these fibers (**Supplemental Figure 3**). In addition to Ppary, several other regulators were predicted to play a role in the transcriptional responses caused by one or more of the fibers (**Table 2**). Notably, the histone demethylase KDM5B was exclusively predicted to control the RS-modulated gene expression profiles, which may at least in part explain the specific gene expression profile induced by RS. Taken together, our analyses identified the nuclear receptor Ppary as potential key upstream regulator mediating the gene expression response to fibers in colonic epithelia.

Modulation of luminal microbiota composition by dietary fibers

Intestinal content of the colon from four mice per dietary treatment was subjected to microbiota quantification and composition analysis. Although all fibers seemed to increase colonic microbiota density compared to the control diet, no statistical significance was reached (**Supplemental Figure 4**). MITChip analysis revealed that all five fibers changed the colonic microbiota composition, except for a single mouse from the RS group that clustered with the mice from the CON group (**Supplemental Figure 5**). The microbial diversity, as determined by the Shannon index, was not significantly different between any of the diets (data not shown). To relate changes in microbiota composition to the different diets, the hybridization signals of the 96 genus-level phylogenetic groups were subjected to redundancy analysis (RDA). Overall, 73.7% of the total variation in microbiota composition was captured within the first two canonical differentiation axes, with diet explaining 34.8% (**Figure 2**). Samples from the RS and control diet clustered separately from IN, AX, FOS and GG. The genus-like groups in the plot that correlated with RS and control diets belong to the *Bacteroidetes* phylum, and also encompassed specific classes of the *Firmicutes* phylum (*Bacilli*, *Clostridium* clusters I, II and IV), and single genus groups of the *Actinobacteria*, *Proteobacteria* and *Deferribacter* phyla. In the opposite direction groups within *Clostridium* cluster XIVa and a specific genus group of the *Bacteroidetes* phylum correlated with IN, GG and FOS. The AX diet-group was positioned centrally in the plot, illustrating that this diet did not clearly correlate with changes in any specific bacterial groups.

Modulation of luminal SCFA concentrations by dietary fibers

Despite differences in fiber type, IN, FOS, GG and AX induced highly similar epithelial cell gene expression responses. Therefore, it was assumed that this might be explained by similar production of fermentation products of these fibers by the microbiota. As the main fermentation metabolites of dietary fibers, SCFA were analyzed in the luminal content. Total SCFA concentrations significantly increased in colonic luminal samples obtained from mice that were fed IN, AX, and GG ($P < 0.05$), while a similar trend was observed for FOS ($P = 0.07$). The highest cumulative SCFA concentrations were observed for mice fed GG, followed by IN. In contrast, the SCFA concentrations in samples obtained from mice fed RS were comparable to those obtained from the control fed animals (**Figure 3A**). In all diet groups, the acetate concentration was highest. Analysis of variance revealed that mean concentrations of both acetate and propionate were significantly different between any of the fibers ($P < 0.05$) (**Figure 3B**). Specifically, acetate concentrations in samples obtained from mice fed AX, IN and GG were significantly higher compared to samples obtained from mice fed CON and RS diet, while propionate concentrations were significantly higher in samples from mice fed FOS, AX, IN and GG compared to CON. Butyrate concentrations were not significantly different between the diet groups but showed a trend ($P = 0.061$). No significant difference was observed for iso-butyrate, valerate or iso-valerate.

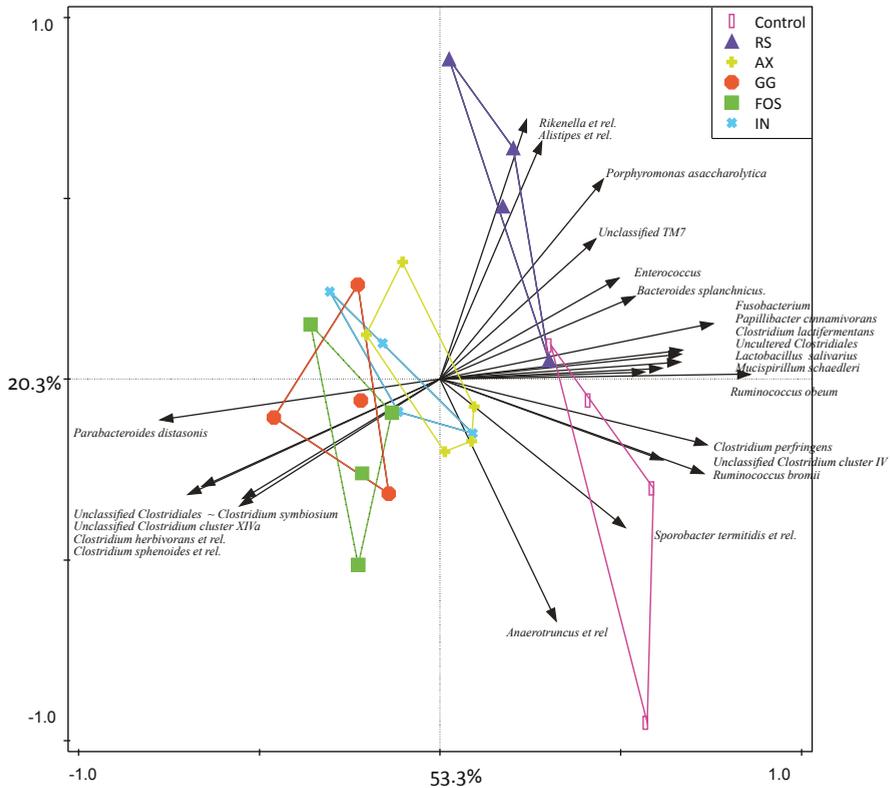


Figure 2 Differential modulation of colonic microbiota composition by dietary fibers

Correlation triplot based on a redundancy analysis (RDA) depicting the relationship between colonic luminal microbiota composition and the differences induced by dietary fibers. Dietary fiber, used as explanatory variable, explained 34.8% of the total variation in the microbiota composition, and 73.7% of that variation was explained by the first two canonical axes shown here. Samples are labelled per diet group, and bacterial groups are indicated by arrows. The arrows point in the direction of maximal variation in the species abundances, and their lengths are proportional to their maximal rate of change. Long arrows correspond to species contributing more to the data set variation. Right-angle projection of a sample dot on a species arrow gives approximate species abundance in the sample.

Integrative analysis of changes in colonic epithelial cell gene expression and luminal microbiota composition

To get further insight in the interaction between changes in gene expression and microbiota composition, which may allow the generation of new hypotheses about potential mechanisms that explain host transcriptional responses to fiber fermentation by

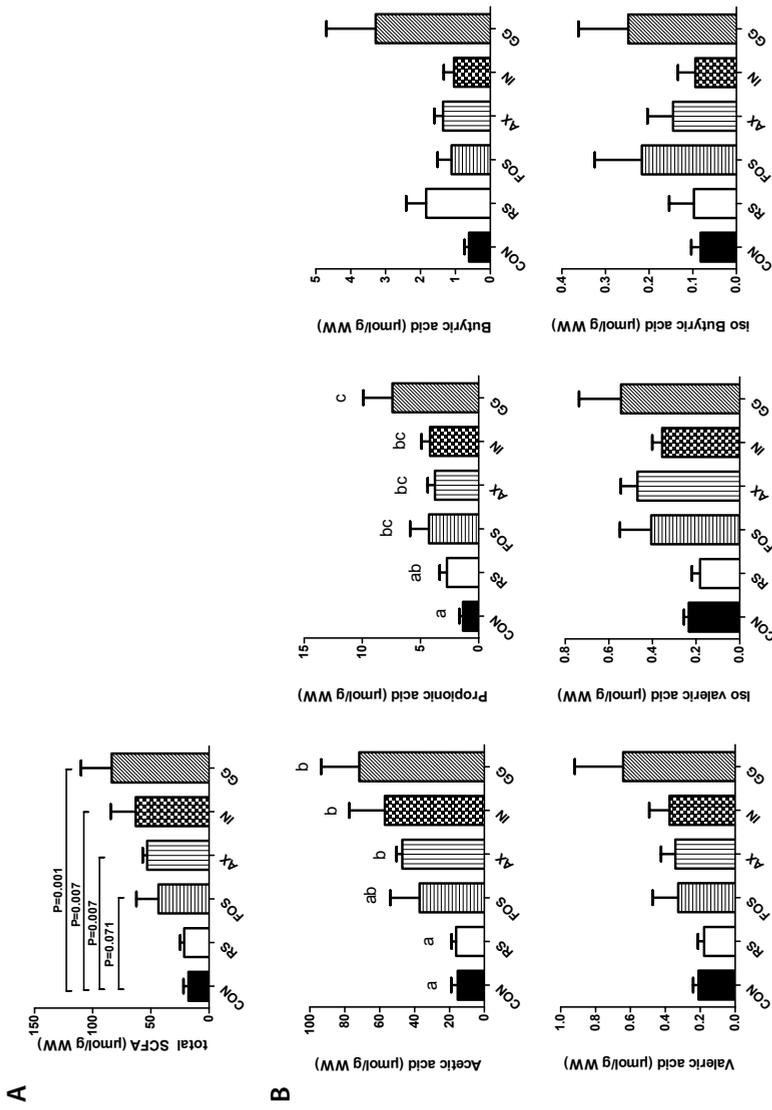


Figure 3 SCFA concentration in colonic luminal samples of mice fed different fibers

Colonic luminal SCFA concentrations in μmol/g colonic content were measured with gas chromatography. The mean ± SEM for **A**) total SCFA concentrations and **B**) individual SCFA for each diet group is represented in a bar plot. Different letters indicate statistical difference between the diet groups as tested with ANOVA and Tukey post hoc test.



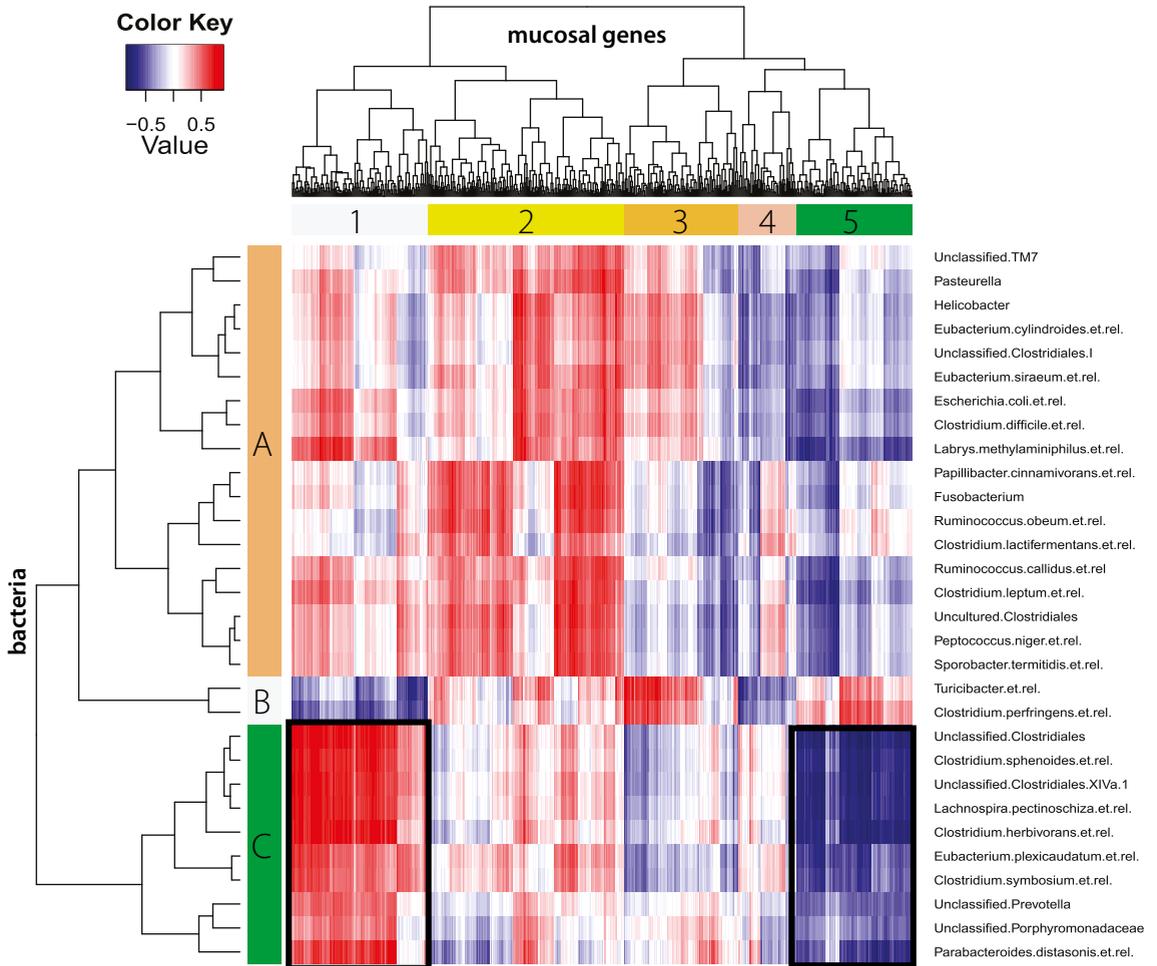


Figure 4 Integration of epithelial cell gene expression with luminal microbiota composition

Sparse PLS canonical correlation analysis was performed to integrate gene expression values with relative abundance data of bacteria for individual mice. The heatmap represents the correlation structure of both dataset; red: positively correlated, blue: negatively correlated. The more intense the color is, the higher the correlation value. Correlation values were subjected to unsupervised hierarchical clustering based on Euclidean distance for both genes and microbial groups. Five main gene clusters (numbers), and three main bacterial clusters (capital letters) were identified.

the microbiota, changes in microbiota composition correlated with changes in colonic epithelial cell gene expression. The correlation pattern between microbiota and gene expression across 22 samples was visualized in a clustered image map. The correlation analysis revealed five clusters of genes and three main clusters of bacterial groups (**Figure 4**). The strongest correlations were found for bacteria in cluster C, which positively correlated with genes in cluster 1, but negatively correlated with genes in cluster 5. This cluster C contained known butyrate-producing bacteria belonging to *Clostridium* cluster XIVa. These bacteria co-clustered with saccharolytic bacteria such as *Parabacteroides distasonis* [42]. Genes in cluster 1 were involved in metabolic, energy-generating and oxidative processes, whereas genes in cluster 5 were involved in adhesion dynamics and signalling. While these processes correlated positively with bacteria from *Clostridium* cluster XIVa, *Turicibacter* et. rel and *Clostridium perfringens* showed negative correlation with these sets of genes. Correlation of the three main SCFA, acetate, propionate and butyrate with host gene expression showed strongest correlation for acetate and propionate (data not shown). Thus, multivariate analyses revealed strong correlations between gene expression changes and relative abundance of bacteria, among which members of *Clostridium* cluster XIVa stood out because of strong association with mucosal gene expression patterns.

Saccharide profiles of the colonic luminal content

Except for RS, the fibers used in this study are known as water-soluble fibres. Hence, to obtain an indication on the extent of degradation of the fibers, the water-soluble fraction of the colonic luminal content was subjected to analysis using HPAEC-PAD. The results revealed that most oligosaccharides from IN, AX, FOS and GG were absent in the colon. The estimated amounts of mono/disaccharides and oligosaccharides are presented in **Figure 5**. No maltodextrins were detected in the samples from IN, AX, FOS and GG, indicating that the corn starch in the diet was completely digested before it reached the colon. As an example, the oligosaccharide profiles of the colonic luminal content of FOS-fed mice compared with FOS and maltodextrins are presented in **Supplemental Figure 6**. Results from the other three fibres were similar to that of FOS. Additionally, size exclusion chromatography indicated that also no water-soluble polysaccharides were present in the case of IN, AX and GG (data not shown), suggesting that like FOS, these fibers were also extensively fermented. In contrast to IN, AX, FOS and GG, maltodextrin was present in the samples from CON and RS (**Figure 5**). The presence of maltodextrins in CON indicated that not all of the corn starch in CON diet was digested and absorbed before reaching the colon. For RS, large amounts of glucose and maltodextrins were present in the colon, showing that RS was degraded by the microbiota to maltodextrins, but also that it was not completely fermented in the colon. In this experiment, however, it was not possible to estimate the rate of dietary fiber degradation in the colon.

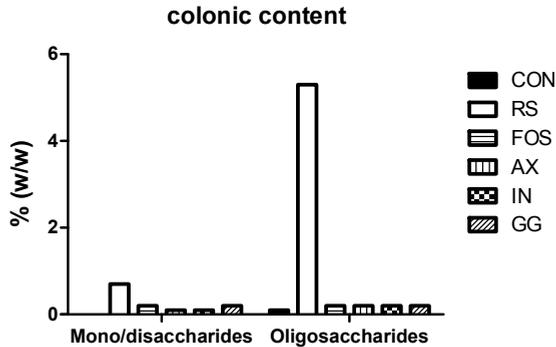


Figure 5 Saccharide content in colonic luminal samples of mice fed different fibers

Saccharide content was analyzed in luminal content of mice (n=1; per group) using HPAEC-PAD. Mono/disaccharides fraction represent glucose, fructose and sucrose, whereas the oligosaccharides fraction represent maltodextrin. Concentrations are expressed as percentage w/w. Quantification of mono-/di-saccharides were based on glucose, whereas quantification of oligosaccharides were based on oligofructose (FOS), with the assumption that FOS is completely soluble and analyzed by HPAEC. The concentration of FOS was corrected for the glucose, fructose and sucrose present.

Discussion

In the present study five different dietary fibers were fed to mice to comprehensively study their effects on epithelial cell gene expression, luminal microbiota composition and SCFA concentrations. Despite differences in source and composition, we found that 4 of the 5 fibers (IN, FOS, AX, and GG) induced a highly similar gene expression and microbiota composition profile in colon, which coincided with increased SCFA concentrations. Another fiber, RS, induced a distinct response and in some cases showed more similar responses to CON, which coincided with the presence of maltodextrin in the luminal content of mice fed with both diets. In addition, the nuclear receptor Ppar γ was identified as a potential key upstream regulator.

Biological implication of gene regulation in epithelial cells

Dietary fibers have diverse effects on the microbiota and host metabolism. To the best of our knowledge, however, it has not yet been determined how in the same experiment, i.e. in mice from the same breeding colony housed in the same facility and fed the same background diet, different dietary fibers modulate colonic gene expression. At the start of this study, we anticipated that fibers that varied in source and composition would all induce a distinct responses in terms of colonic epithelial gene expression. However, our results showed that the diet groups could be grouped in 3 main clusters (CON, RS and the

other 4 fibers [IN, FOS, AX, and GG]), and that many of the transcriptional responses for these latter 4 fibers were conserved on the level of biological processes, as determined by gene set enrichment analysis. However, specific genes were more variable.

It is assumed that the effects of fibers are mainly mediated by the type and level of SCFA that are produced during their fermentation by the microbiota. These SCFA have been shown to elicit diverse effects on colonic gene expression patterns [43-45]. Our results support this view, since the separation of the epithelial gene expression profiles in 3 clusters coincided with the cumulative SCFA levels measured in the colonic lumen. As a consequence, the biological implications of gene regulation induced by the fibers that yielded increased SCFA concentrations are very much the same. These fibers all induced genes involved in the glycolysis, TCA cycle, electron transport chain and oxidative phosphorylation, all pathways responsible for generating energy by substrate oxidation. This is to be expected because SCFAs, and especially butyrate, are known to serve as preferred energy source for colonocytes as a precursor to the TCA cycle and electron transport chain [46,47]. Moreover, facilitation by the microbiota of oxidative metabolism of glucose in colonocytes has been reported [48]. Because the generation of ATP results in the formation of reactive oxygen species [49], this may explain the activation of NRF2-controlled antioxidant response pathways [50], that comprise induction of GST levels, as also has been observed by others [51]. Large clusters of gene sets suppressed by fibers represented processes describing gene transcription and translation, which is indicative for reduced epithelial cell proliferation, which is in agreement with findings that fibers inhibit growth and proliferation (reviewed in [52]). Since this also was observed for RS, that did not increase SCFA concentrations, our results suggest that at least part of this effect is independent of SCFA, as has been suggested [52]. We hypothesize that the observed fiber-specific effects are also mediated by specific degradation products and not SCFA [8]. Preliminary analyses of mono-, di- and oligosaccharides in the intestinal lumen samples revealed that contrary to mice fed the other diets, CON and RS fed mice contained maltodextrin in their colonic lumen, which might explain the distinct response (Fig. 5).

Ppar γ is identified as central regulator of transcriptional responses to fermented fibers

Our analysis consistently identified Ppar γ as likely regulator that mediate the effects of the fermented fibers on gene expression. In line with this observation, target genes of the nuclear receptor Ppar were found to be consistently induced, albeit to a different extent, for all of these 4 fibers. In addition, it has been reported that in proximal colon Ppar γ mainly regulates genes involved in energy metabolism, in particular lipid catabolism, but also affects signaling, motility and cell adhesion [53]. In our study, especially the IN fed mice displayed an increased expression level of genes involved in lipid metabolic processes. Since Ppar γ has recently been shown to be activated by SCFAs, in particular butyrate [54], this provides strong evidence for a molecular mechanism by which

fermentable fibers may modulate colonic gene expression. Nevertheless, the variability of gene response magnitude as well as specific gene expression patterns that may be related to Ppar γ activation and the role of SCFA production patterns warrant further research. Activation of PPAR γ is also of interest because this transcription factor has been demonstrated to coordinate the expression of anti-inflammatory properties in inflammatory bowel disease [55] and appears to play a pivotal role in the interplay between metabolism and immune function regulation [56].

Abundance of Clostridium cluster XIVa correlates with epithelial cell metabolic pathways

We showed that the fibers yielding increased SCFA concentrations in the colonic lumen all increased the transcription of genes involved in metabolic processes associated with energy metabolism. Our correlation analysis demonstrated multiple relationships between the microbiota composition and these changing host gene expression patterns. In particular bacterial groups within *Clostridium* cluster XIVa positively correlated with genes involved in energy metabolism. Next to primary degrading bacterial species, this bacterial group is known to encompass many secondary fermenters, of which several have been shown to produce butyrate as their metabolic end product [57]. Unfortunately, and analogous to many other studies, the data presented here are based on single time-point measurements, and thereby fail to represent actual production or absorption rates of SCFA. Such flux data could give a considerable refinement to our understanding of the rate of production of butyrate by these bacteria and the actual levels of butyrate flux experienced by the colonic epithelia, respectively [58].

Conclusion

Our results provide a comprehensive overview on the effects of five fibers in the murine colon, which suggest that despite different source and composition, fermentable fibers are inducing a highly similar mucosal response that may at least be partially governed by Ppar γ . –

Acknowledgments

We would like to thank Dr Diederick Meyer (Sensus) and Dr Hans van der Saag (BioActor) for their kind gifts of inulin and oligofructose, respectively arabinoxylan.

This work was (co)financed by the Netherlands Consortium for Systems Biology (NCSB), which is part of the Netherlands Genomics Initiative / Netherlands Organization for Scientific Research.

The authors have declared no conflict of interest.

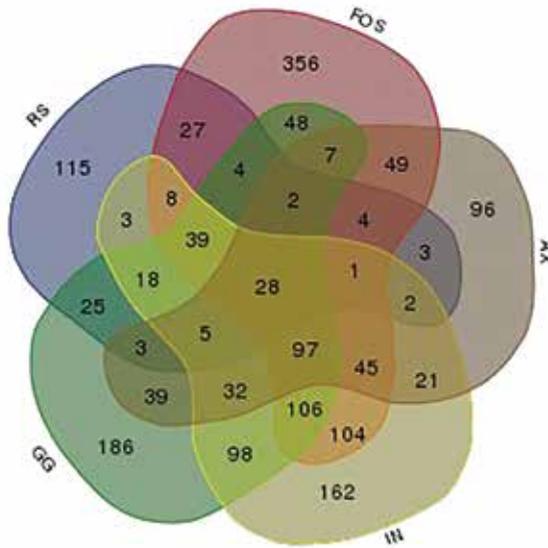
References

1. Phillips GO, Cui SW (2011) An introduction: Evolution and finalisation of the regulatory definition of dietary fibre. *Food Hydrocolloid* 25: 139-143.
2. Jones JM (2014) CODEX-aligned dietary fiber definitions help to bridge the 'fiber gap'. *Nutr J* 13: 34.
3. Aune D, Chan DS, Lau R, Vieira R, Greenwood DC, et al. (2011) Dietary fibre, whole grains, and risk of colorectal cancer: systematic review and dose-response meta-analysis of prospective studies. *BMJ* 343: d6617.
4. Balakrishnan M, Floch MH (2012) Prebiotics, probiotics and digestive health. *Curr Opin Clin Nutr Metab Care* 15: 580-585.
5. Slavin JL (2013) Fiber and prebiotics: mechanisms and health benefits. *Nutrients* 5: 1417-1435.
6. Eswaran S, Muir J, Chey WD (2013) Fiber and functional gastrointestinal disorders. *Am J Gastroenterol* 108: 718-727.
7. Nyangale EP, Mottram DS, Gibson GR (2012) Gut microbial activity, implications for health and disease: the potential role of metabolite analysis. *J Proteome Res* 11: 5573-5585.
8. Russell WR, Hoyles L, Flint HJ, Dumas ME (2013) Colonic bacterial metabolites and human health. *Curr Opin Microbiol* 16: 246-254.
9. Flint HJ, Bayer EA, Rincon MT, Lamed R, White BA (2008) Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nat Rev Microbiol* 6: 121-131.
10. Van den Abbeele P, Gerard P, Rabot S, Bruneau A, El Aidy S, et al. (2011) Arabinoxylans and inulin differentially modulate the mucosal and luminal gut microbiota and mucin-degradation in humanized rats. *Environ Microbiol* 13: 2667-2680.
11. Van den Abbeele P, Venema K, Van de Wiele T, Verstraete W, Possemiers S (2013) Different human gut models reveal the distinct fermentation patterns of Arabinoxylan versus inulin. *J Agric Food Chem* 61: 9819-9827.
12. El Kaoutari A, Armougom F, Gordon JI, Raoult D, Henricsson B (2013) The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nat Rev Microbiol* 11: 497-504.
13. Hooper LV, Midtvedt T, Gordon JI (2002) How host-microbial interactions shape the nutrient environment of the mammalian intestine. *Annu Rev Nutr* 22: 283-307.
14. Barboza M, Sela DA, Pirim C, Locascio RG, Freeman SL, et al. (2009) Glycoprofiling bifidobacterial consumption of galacto-oligosaccharides by mass spectrometry reveals strain-specific, preferential consumption of glycans. *Appl Environ Microbiol* 75: 7319-7325.
15. Riviere A, Moens F, Selak M, Maes D, Weckx S, et al. (2014) The ability of bifidobacteria to degrade arabinoxylan oligosaccharide constituents and derived oligosaccharides is strain dependent. *Appl Environ Microbiol* 80: 204-217.
16. Müller M, Kersten S (2003) Nutrigenomics: goals and strategies. *Nat Rev Genet* 4: 315-322.
17. Afman L, Muller M (2006) Nutrigenomics: from molecular nutrition to prevention of disease. *J Am Diet Assoc* 106: 569-576.
18. Keenan MJ, Martin RJ, Raggio AM, McCutcheon KL, Brown IL, et al. (2012) High-amylose resistant starch increases hormones and improves structure and function of the gastrointestinal tract: a microarray study. *J Nutrigenet Nutrigenomics* 5: 26-44.
19. Haenen D, Souza da Silva C, Zhang J, Koopmans SJ, Bosch G, et al. (2013) Resistant starch induces catabolic but suppresses immune and cell division pathways and changes the microbiome in the proximal colon of male pigs. *J Nutr* 143: 1889-1898.
20. Rodenburg W, Keijer J, Kramer E, Vink C, van der Meer R, et al. (2008) Impaired barrier function by dietary fructo-oligosaccharides (FOS) in rats is accompanied by increased colonic mitochondrial gene expression. *BMC Genomics* 9: 144.
21. Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R, et al. (2009) The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Sci Transl Med* 1: 6ra14.
22. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5: R80.
23. Lin K, Kools H, de Groot PJ, Gavai AK, Basnet RK, et al. (2011) MADMAX - Management and analysis database for multiple ~omics experiments. *J Integr Bioinform* 8: 160.

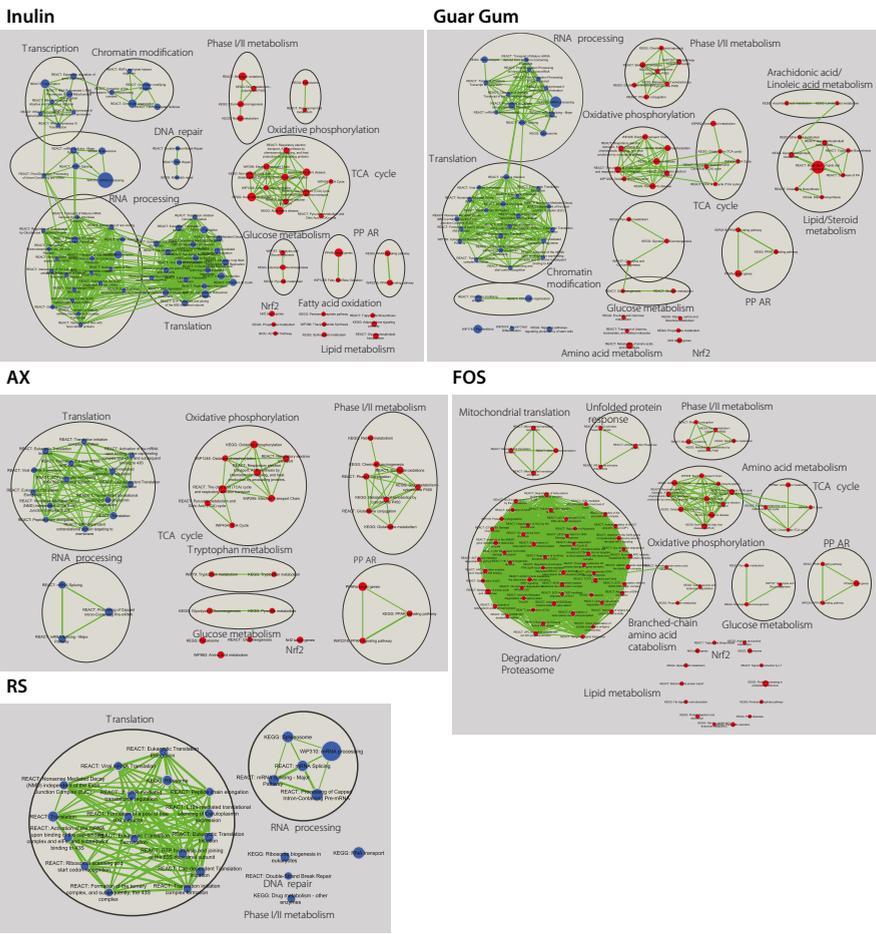
24. Piccolo SR, Withers MR, Francis OE, Bild AH, Johnson WE (2013) Multiplatform single-sample estimates of transcriptional activation. *Proc Natl Acad Sci U S A* 110: 17778-17783.
25. Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3: Article3.
26. Sartor MA, Tomlinson CR, Wesselkamper SC, Sivaganesan S, Leikauf GD, et al. (2006) Intensity-based hierarchical Bayes method improves testing for differentially expressed genes in microarray experiments. *BMC Bioinformatics* 7: 538.
27. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102: 15545-15550.
28. Allison DB, Cui X, Page GP, Sabripour M (2006) Microarray data analysis: from disarray to consolidation and consensus. *Nat Rev Genet* 7: 55-65.
29. Abatangelo L, Maglietta R, Distaso A, D'Addabbo A, Creanza TM, et al. (2009) Comparative study of gene set enrichment methods. *BMC Bioinformatics* 10: 275.
30. Merico D, Isserlin R, Stueker O, Emili A, Bader GD (2010) Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One* 5: e13984.
31. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, et al. (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* 14: 128.
32. Yu Z, Morrison M (2004) Improved extraction of PCR-quality community DNA from digesta and fecal samples. *Biotechniques* 36: 808-812.
33. Geurts L, Lazarevic V, Derrien M, Everard A, Van Roye M, et al. (2011) Altered gut microbiota and endocannabinoid system tone in obese and diabetic leptin-resistant mice: impact on apelin regulation in adipose tissue. *Front Microbiol* 2: 149.
34. Lahti L, Torrente A, Elo LL, Brazma A, Rung J (2013) A fully scalable online pre-processing algorithm for short oligonucleotide microarray atlases. *Nucleic Acids Res* 41: e110.
35. Lepš J, Šmilauer P (2003) *Multivariate Analysis of Ecological Data using CANOCO*. Cambridge, UK: Cambridge University Press. 282 p.
36. Haenen D, Zhang J, Souza da Silva C, Bosch G, van der Meer IM, et al. (2013) A diet high in resistant starch modulates microbiota composition, SCFA concentrations, and gene expression in pig intestine. *J Nutr* 143: 274-283.
37. Jonathan MC, van den Borne JJGC, van Wiechen P, Souza da Silva C, Schols HA, et al. (2012) In vitro fermentation of 12 dietary fibres by faecal inoculum from pigs and humans. *Food Chem* 133: 889-897.
38. Boulesteix AL, Strimmer K (2007) Partial least squares: a versatile tool for the analysis of high-dimensional genomic data. *Brief Bioinform* 8: 32-44.
39. Le Cao KA, Martin PG, Robert-Granie C, Besse P (2009) Sparse canonical methods for biological data integration: application to a cross-platform study. *BMC Bioinformatics* 10: 34.
40. Gonzalez I, Cao KA, Davis MJ, Dejean S (2012) Visualising associations between paired 'omics' data sets. *BioData Min* 5: 19.
41. Le Cao KA, Gonzalez I, Dejean S (2009) integrOmics: an R package to unravel relationships between two omics datasets. *Bioinformatics* 25: 2855-2856.
42. Sakamoto M, Benno Y (2006) Reclassification of *Bacteroides distasonis*, *Bacteroides goldsteinii* and *Bacteroides merdae* as *Parabacteroides distasonis* gen. nov., comb. nov., *Parabacteroides goldsteinii* comb. nov. and *Parabacteroides merdae* comb. nov. *Int J Syst Evol Microbiol* 56: 1599-1605.
43. Topping DL, Clifton PM (2001) Short-chain fatty acids and human colonic function: roles of resistant starch and nonstarch polysaccharides. *Physiol Rev* 81: 1031-1064.
44. Vanhoutvin SA, Troost FJ, Hamer HM, Lindsey PJ, Koek GH, et al. (2009) Butyrate-induced transcriptional changes in human colonic mucosa. *PLoS One* 4: e6759.
45. Fung KY, Cosgrove L, Lockett T, Head R, Topping DL (2012) A review of the potential mechanisms for the lowering of colorectal oncogenesis by butyrate. *Br J Nutr* 108: 820-831.
46. Roediger WE (1982) Utilization of nutrients by isolated epithelial cells of the rat colon. *Gastroenterology* 83: 424-429.

47. Donohoe DR, Garge N, Zhang X, Sun W, O'Connell TM, et al. (2011) The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon. *Cell Metab* 13: 517-526.
48. Donohoe DR, Wali A, Brylawski BP, Bultman SJ (2012) Microbial regulation of glucose metabolism and cell-cycle progression in mammalian colonocytes. *PLoS One* 7: e46589.
49. Davies KJ (1995) Oxidative stress: the paradox of aerobic life. *Biochem Soc Symp* 61: 1-31.
50. Klaassen CD, Reisman SA (2010) Nrf2 the rescue: effects of the antioxidative/electrophilic response on the liver. *Toxicol Appl Pharmacol* 244: 57-65.
51. Yaku K, Enami Y, Kurajyo C, Matsui-Yuasa I, Konishi Y, et al. (2012) The enhancement of phase 2 enzyme activities by sodium butyrate in normal intestinal epithelial cells is associated with Nrf2 and p53. *Mol Cell Biochem* 370: 7-14.
52. Zeng H, Lazarova DL, Bordonaro M (2014) Mechanisms linking dietary fiber, gut microbiota and colon cancer prevention. *World J Gastrointest Oncol* 6: 41-51.
53. Su W, Bush CR, Necela BM, Calcagno SR, Murray NR, et al. (2007) Differential expression, distribution, and function of PPAR-gamma in the proximal and distal colon. *Physiol Genomics* 30: 342-353.
54. Alex S, Lange K, Amolo T, Grinstead JS, Haakonsson AK, et al. (2013) Short-chain fatty acids stimulate angiopoietin-like 4 synthesis in human colon adenocarcinoma cells by activating peroxisome proliferator-activated receptor gamma. *Mol Cell Biol* 33: 1303-1316.
55. Dubuquoy L, Rousseaux C, Thuru X, Peyrin-Biroulet L, Romano O, et al. (2006) PPARgamma as a new therapeutic target in inflammatory bowel diseases. *Gut* 55: 1341-1349.
56. Hou Y, Moreau F, Chadee K (2012) PPARgamma is an E3 ligase that induces the degradation of NFkappaB/p65. *Nat Commun* 3: 1300.
57. Louis P, Flint HJ (2009) Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. *FEMS Microbiol Lett* 294: 1-8.
58. den Besten G, Havinga R, Bleeker A, Rao S, Gerding A, et al. (2014) The short-chain fatty acid uptake fluxes by mice on a guar gum supplemented diet associate with amelioration of major biomarkers of the metabolic syndrome. *PLoS One* 9: e107392.
59. Heber S, Sick B (2006) Quality assessment of Affymetrix GeneChip data. *OMICS* 10: 358-368.
60. Dai M, Wang P, Boyd AD, Kostov G, Athey B, et al. (2005) Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res* 33: e175.
61. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, et al. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4: 249-264.
62. Rajilic-Stojanovic M, Heilig HG, Molenaar D, Kajander K, Surakka A, et al. (2009) Development and application of the human intestinal tract chip, a phylogenetic microarray: analysis of universally conserved phylotypes in the abundant microbiota of young and elderly adults. *Environ Microbiol* 11: 1736-1751.
63. Reikvam DH, Derrien M, Islam R, Erofeev A, Grcic V, et al. (2012) Epithelial-microbial crosstalk in polymeric Ig receptor deficient mice. *Eur J Immunol* 42: 2959-2970.
64. Lahti L, Salojärvi J (2014) The microbiome R package: standard routines for the preprocessing and analysis of microbiomics profiling data. <http://microbiomegithub.com>.
65. Suzuki MT, Taylor LT, DeLong EF (2000) Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5'-nuclease assays. *Appl Environ Microbiol* 66: 4605-4614.
66. van den Bogert B, de Vos WM, Zoetendal EG, Kleerebezem M (2011) Microarray analysis and barcoded pyrosequencing provide consistent microbial profiles depending on the source of human intestinal samples. *Appl Environ Microbiol* 77: 2071-2080.

Supplemental Information

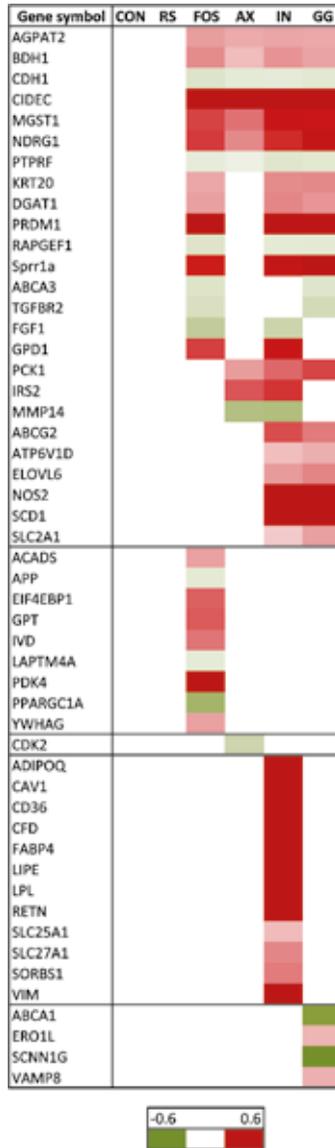


Supplemental Figure 1 Overlap of genes significantly regulated by each dietary fiber ($P < 0.01$)



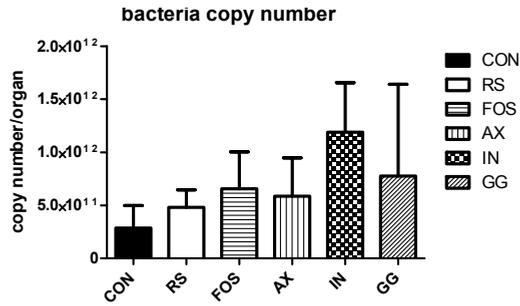
Supplemental Figure 2

Enrichment Map of gene sets that were changed by fiber compared to control. GSEA was performed to identify functional gene sets, i.e. metabolic pathways or signaling transduction routes, that were changed by each fiber compared to control ($p < 0.001$, $FDR < 0.01$). Nodes represent functional gene sets, and edges between nodes represent their similarity. A red node indicates induction of a gene set, a blue node indicates suppression of a gene set, and a white node indicates no significant regulation of a gene set by a fiber compared to control. Node size represents the gene set size, and edge thickness represents the degree of overlap between 2 connected gene sets. Clusters were grouped and manually labeled to highlight the prevalent biologic functions among related gene sets.



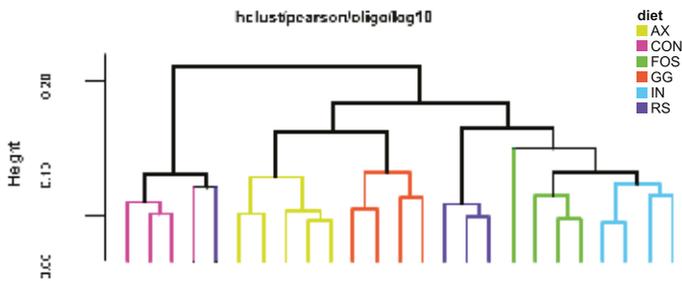
Supplemental Figure 3 Upstream regulator analysis

Pparγ target genes were determined by Ingenuity Pathway Analysis. A heatmap represents the relative gene expression values for each fiber diet compared to control. Red indicates increased expression, while green indicates decreased expression. The relative gene expression values were log₂ transformed, i.e. a fold change of 0.6 means 1.5 fold increase, while -0.6 means -1.5 decrease in expression by the fiber compared to control.



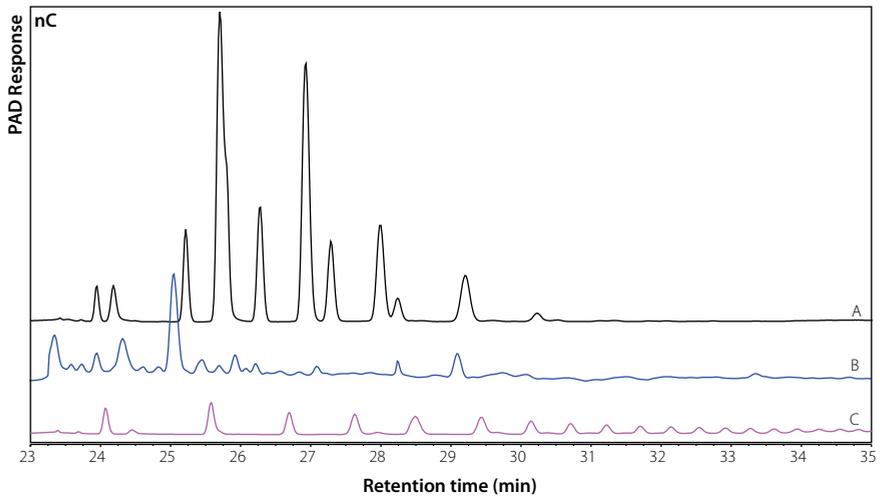
Supplemental Figure 4 Quantitative PCR on total bacteria

16S rRNA gene-targeted qPCR was used to assess total bacterial numbers. The copy number per 16S rRNA gene was calculated back to total copy number per organ weight.



Supplemental Figure 5 Clustering of MITChip profiles at the probe-level

Pearson distance-based clustering of the samples on log₁₀ transformed probe level data of the MITChip.



Supplemental Figure 6 Representative monomer and oligomer profiles analyzed using HPAEC-PAD

Oligosaccharide profiles of (A) oligofructose (FOS), (B) colonic luminal content of a FOS-fed mouse, as analysed by high performance anion exchange chromatography coupled to pulsed amperometric detection (HPAEC-PAD). The profile of (C) maltodextrin is presented as a comparison.



Supplemental Table 1 Diet composition

Based on formula #	1. Low fat - palm oil - CON		2. Low fat - palm oil - IN		3. Low fat - palm oil - FOS	
	D12450B		10% INULIN		10% FOS	
Ingredient	gm	% (w/w)	gm	% (w/w)	gm	% (w/w)
Casein, lactic	200.00	18.96	200.00	18.96	200.00	18.96
L-Cystine	3.00	0.28	3.00	0.28	3.00	0.28
Corn Starch (control fiber)	427.20	40.49	321.70	30.49	321.70	30.49
Maltodextrin	100.00	9.48	100.00	9.48	100.00	9.48
Sucrose	172.80	16.38	172.80	16.38	172.80	16.38
Experimental fiber	n.a.		105.50	10.00	105.50	10.00
Cellulose, BW200	50.00	4.74	50.00	4.74	50.00	4.74
Soybean Oil	25.00	2.37	25.00	2.37	25.00	2.37
Palm oil	20.00	1.90	20.00	1.90	20.00	1.90
Mineral Mix S10026*	10.00	0.95	10.00	0.95	10.00	0.95
DiCalcium Phosphate	13.00	1.23	13.00	1.23	13.00	1.23
Calcium Carbonate	5.50	0.52	5.50	0.52	5.50	0.52
Potassium Citrate, 1 H ₂ O	16.50	1.56	16.50	1.56	16.50	1.56
Vitamin Mix V10001**	10.00	0.95	10.00	0.95	10.00	0.95
Choline Bitartrate	2.00	0.19	2.00	0.19	2.00	0.19
FD&C Yellow Dye #5	0.05	0.00	0.05	0.00	0.05	0.00
FD&C Red Dye #40	0.00	0.00	0.00	0.00	0.00	0.00
FD&C Blue Dye #1	0.00	0.00	0.00	0.00	0.00	0.00
Total	1055.05	100.00	1055.05	100.00	1055.05	100.00

*Mineral mix S10026 contains the following (g/kg mineral mix): 41.9 magnesium oxide, 257.6 magnesium sulfate·7H₂O, 259 sodium chloride, 1.925 chromium KSO₄·12H₂O, 1.05 cupric carbonate, 0.035 potassium iodate, 21 ferric citrate, 12.25 manganous carbonate, 0.035 sodium selenite, 5.6 zinc carbonate, 0.20 sodium fluoride, 0.30 ammonium molybdate·4H₂O, 399.105 sucrose.

**Vitamin mix V10001 contains the following (g/kg vitamin mix): 0.80 retinyl palmitate, 1.0 cholecalciferol, 10 all-rac-a-tocopheryl acetate, 0.08 menadione sodiumbisulfite, 2.0 biotin (1.0%), 1.0 cyanocobalamin (0.1%), 0.20 folic acid, 3.0 nicotinic acid, 1.6 calcium pantothenate, 0.70 pyridoxine-HCl, 0.60 riboflavin, 0.60 thiamin-HCl, and 978.42 sucrose.

4. Low fat - palm oil - AX		5. Low fat - palm oil - GG		6. Low fat - palm oil - RS	
10% NAXUS		10% GUAR GUM		20% Resistant Starch	
gm	% (w/w)	gm	% (w/w)	gm	% (w/w)
200.00	18.96	200.00	18.96	200.00	18.96
3.00	0.28	3.00	0.28	3.00	0.28
321.70	30.49	321.70	30.49	216.20	20.49
100.00	9.48	100.00	9.48	100.00	9.48
172.80	16.38	172.80	16.38	172.80	16.38
105.50	10.00	105.50	10.00	211.00	20.00
50.00	4.74	50.00	4.74	50.00	4.74
25.00	2.37	25.00	2.37	25.00	2.37
20.00	1.90	20.00	1.90	20.00	1.90
10.00	0.95	10.00	0.95	10.00	0.95
13.00	1.23	13.00	1.23	13.00	1.23
5.50	0.52	5.50	0.52	5.50	0.52
16.50	1.56	16.50	1.56	16.50	1.56
10.00	0.95	10.00	0.95	10.00	0.95
2.00	0.19	2.00	0.19	2.00	0.19
0.05	0.00	0.05	0.00	0.05	0.00
0.00	0.00	0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00	0.00	0.00
1055.05	100.00	1055.05	100.00	1055.05	

Supplemental Methods

Microarray processing and data analysis

Total RNA (100 ng) was used for whole transcript cDNA synthesis by using the Ambion WT expression kit (Life Technologies, Bleiswijk, the Netherlands) and subsequently labelled using the Affymetrix GeneChip WT Terminal Labeling Kit. Samples were hybridized on Mouse Gene 1.1 ST arrays (Affymetrix, Santa Clara, US), washed, stained, and scanned on an Affymetrix GeneTitan instrument. Detailed protocols for array handling can be found in the GeneChip WT Terminal Labeling and Hybridization User Manual (P/N 702808, Rev. 7; Affymetrix). Packages from the Bioconductor project [22], integrated in an online pipeline [23], were used to analyze the array data. Various advanced-quality metrics, diagnostic plots, pseudoimages, and classification methods were used to determine the quality of the arrays before statistical analysis [59]. The probes on the Mouse Gene 1.1 ST array were redefined using current genome information [60]. In this study, probes were reorganized on the basis of the gene definitions available in the NCBI Mus musculus Entrez Gene database based on the mouse genome build 38 patch release 1 (GRCm38.p1) (custom CDF v17). Normalized gene expression estimates were obtained from the raw intensity values using the robust multiarray analysis (RMA) preprocessing algorithm available in the library 'AffyPLM' using default settings [61]. Differentially expressed probe sets (genes) were identified by using linear models, applying moderated t-statistics that implemented empirical Bayes regularization of standard errors [25]. To adjust for both the degree of independence of variances relative to the degree of identity and the relation between variance and signal intensity, the moderated t-statistic was extended by a Bayesian hierarchical model to define an intensity-based moderated t-statistic [26].

Microbiota analysis

MITChip

This phylogenetic microarray was designed using criteria of the Human Intestinal Tract Chip (HITChip) [62]. The MITChip consists of 3,580 different oligonucleotides specific for the mouse intestinal microbiota [33,62,63]. The array targets the V1 and V6 regions of bacterial 16S rRNA genes. The 16S rRNA genes were amplified from twenty nanogram of intestinal extracted DNA with the primers T7prom-Bact-27-F and Uni-1492-R (see Table). PCR products were then transcribed, and RNA was labelled with Cy3 and Cy5 dyes and fragmented as described previously [33,62,63]. Finally the samples were hybridized on the arrays at 62.5°C for 16 hours in a rotation oven (Agilent Technologies, Amstelveen, the Netherlands). After washing and scanning of the slides, data was extracted with the Agilent Feature Extraction software, version 9.1. The data was normalized and analysed using a set of R-based scripts in combination with a custom-designed relational database [64], which operates under the MySQL database management system.

Quantification of bacterial community

Quantification of the bacterial 16S rRNA gene was performed by a qPCR assay using the primers developed by Suzuki *et al.* [65]. The qPCRs were performed in 384-well plates (BioRad, Veenendaal, the Netherlands) sealed with a film (Microseal B film, Bio-Rad) using a MyIQ cyclor with MyIQ software (version 1.0.410, Bio-Rad). The reactions were carried out in a total volume of 12.5 μ l consisting of 1x IQ SYBR green Supermix (Bio-Rad), 200 nM of the forward and reverse primer and 2 μ l of template DNA, and the cycling program and melting curve analysis as previously described [66]. The standard curve consisting of a 8-fold dilution series was a 16S rRNA gene PCR product of *Escherichia coli* top10.

Table List of primers used for microbiome analyses [62,65]

Primer name	Sequence	Application
T7prom-Bact-27-F	5'-TGA ATT GTA ATA CGA CTC ACT ATA GGG GTT TGA TCC TGG CTC AG-3'	MITChip
Uni-1492-R	5'-CGG CTA CCT TGT TAC GAC-3'	MITChip
PROK1492R	5'-GGW TAC CTT GTT ACG ACT T-3'	qPCR
BACT1369F	5'-CGG TGA ATA CGT TCY CGG-3'	qPCR

Oligosaccharide analysis of colonic luminal content by high performance anion exchange chromatography

Frozen samples of colonic content of mice (n=1 per group) fed different fiber diets were used for analysis of mono-, di- and oligosaccharides. The conditions used for oligosaccharide profiling with high performance anion exchange chromatography coupled to pulsed amperometric detection (HPAEC-PAD) have been described in detail before [37]. In brief, samples were weighed and suspended in water (20 mg/mL), followed by mixing and heating in a boiling water bath for 5 min. After cooling to room temperature, the samples were centrifuged (10,000 \times g, 5 min), and the supernatant (25 μ L) was injected into an ICS 3000 HPAEC system (Dionex Corporation, Sunnyvale, CA, US) with pulsed amperometric detection. One gradient was used for all samples, where monomers were separated from the oligomers. Quantification of mono-/di-saccharides were based on glucose, whereas quantification of oligosaccharides were based on oligofructose (FOS), with the assumption that FOS are completely soluble and analysed by HPAEC. The concentration of FOS was corrected for glucose, fructose and sucrose present.



Chapter 4

Linking the fate of dietary fibers in the murine caecum to microbial transcriptome patterns

Floor Hugenholtz*, Katja Lange*, Mark Davids, Peter Schaap, Michael Müller,
Guido J.E.J. Hooiveld, Michiel Kleerebezem, Hauke Smidt

*These authors contributed equally to this work

In preparation.



Abstract

The microbiota of the gastrointestinal tract plays a key role in the degradation of food components that escape digestion by host enzymes. Complex metabolic networks of interacting microbes in the gastrointestinal tract of humans and other mammals yield a wide range of metabolites of which the short chain fatty acids (SCFA), in particular butyrate, acetate, and propionate, are the most abundant products of carbohydrate fermentation. Here we studied the quantitative interactions between diet, microbiota and host and modelled the multivariate data using Systems Biology approaches.

The experiments targeted the caecum of conventionally raised mice that were fed different fiber-containing diets. Microbiota composition was assessed using phylogenetic microarray technology, and was complemented with metatranscriptome, metabolome and host mucosal tissue transcriptome data. Relative abundance of butyrate producing bacteria correlated with host genes involved in energy metabolism and affecting the immune system. Moreover the metatranscriptome revealed distinct activities of bacterial families in the fermentation of fibers into SCFA. The *Bacteroidaceae*, *Porphyromonadaceae*, *Verrucomicrobioaceae*, *Bifidobacteriaceae*, *Lachnospiraceae*, *Clostridiaceae*, *Eubacteriaceae*, several Bacilli families, *Ruminococcaceae* and *Erysipelotrichaceae* all took part in fiber utilization, expressing genes encoding glycosidases and/or sugar transport systems. All families expressed in different ratios genes that code for enzymes involved in the production of SCFA. Overall different dietary fibers induce distinct changes in the caecal microbiota, their functional activities and SCFA production with profound effects on host metabolic and immune function in the caecum.

Introduction

A diet rich in fiber has a beneficial health effect by promoting gastrointestinal homeostasis, and decreasing risk for obesity, cancer and metabolic disorders. The enzyme repertoire of humans and mammals is commonly not able to catabolize non-starch polysaccharides, like those derived from plant cell walls, resistant starches, and oligosaccharides, whereas a broad range of intestinal bacteria can ferment these dietary fibers. The intestinal microbiota is diverse, dense, metabolically active, and largely saccharolytic [1], [2], [3]. Most members of the large intestinal microbiota are not depending on the availability of simple sugars, but are rather able to derive carbon and energy from the breakdown of a range of complex carbohydrates, sometimes involving single members of the ecosystem, but often requiring a concerted community effort.

The bulk of the fiber fermentation takes place in the large intestine, where the fibers are the main energy source for the microbiota. Dietary fiber is fermented into a range of metabolites, of which the short chain fatty acids (SCFA) acetate, propionate and butyrate are the most abundant [4], [5], [6], [7]. The SCFA production pattern is dependent on diet composition, transit time in the small and large intestine, intestinal pH, and the microbial species and their relative abundance within the microbiota [8], [9], [10]. SCFA are taken up by the mucosa, and have been reported to affect numerous processes related to the immune system [11] and energy metabolism [12]. Additionally, the intestinal microbiota exerts a major influence on host physiology, including the tuning of the host's immune and metabolic state [13]. SCFA are readily absorbed by the colonic epithelium where butyrate serves as the main energy source, although in smaller amounts acetate and propionate can also be metabolised [14]. After absorption and metabolic conversion, the remainder of the SCFA enter the portal blood and are processed by the liver, the central metabolic organ in the body [15]. Intestine-derived acetate is incorporated in fatty acid synthesis in the liver and in part transferred to the peripheral tissue, whereas propionate induces lipogenesis and gluconeogenesis in liver [16].

The amount and ratio of SCFA can be affected by specific types of fiber [5], [17]; **Chapter 3**. The relative abundance of specific groups of bacteria is known to increase when certain types of fibers are consumed. For example, when resistant starch is consumed, the relative abundance of species of *Clostridium* cluster IV that includes butyrate producers is increased, while consumption of FOS and IN tend to expand bifidobacterial populations [18], [19]. The explanation for these type of community responses are rather complex. For example a dietary intervention with resistant starch usually results in increased butyrate concentrations in the intestine, indicative of complex food webs consisting of primary degraders such as those belonging to the *Actinobacteria* (including the *Bifidobacteriaceae*) and Bacteroidetes that ferment starch and produce lactate, succinate, acetate and propionate, which are subsequently taken up by secondary fermenters to produce mainly acetate, propionate and butyrate, explaining the emergence of elevated butyrate levels

upon resistant starch feeding via a cross-feeding (syntrophic) interaction within the microbial consortium [20],[5].

In our previous study mice were fed five different fibers or a control diet containing corn starch that is readily degraded and largely absorbed within the small intestine [Chapter 3]. Colonic samples obtained from the non-starch fiber-fed mice showed an overall increase in luminal SCFA concentrations, which coincided with an increased relative abundance of the *Bacteroides distasonis* group and *Clostridium* cluster XIVa species in the colonic microbiota. These bacteria most likely act as primary or secondary fiber degraders, although data from cultured representatives of the detected species are not available. To further specify the microbial involvement in the degradation of these fibers, and to study the effects on the host mucosal tissue, this study focuses on the main fiber fermentation location in the mouse intestine, i.e. the caecum [Chapter 3]. The large caecal volume is compatible with multiple analyses in individual mice, including the determination of SCFA concentrations, microbiota composition and its activity using metatranscriptome analysis. Furthermore, these multivariate microbial and metabolic datasets were correlated to host mucosal responses in the caecum, determined by tissue transcriptome analyses, aiming to construct models of diet-microbe-host interactions.

Material and methods

Animals, diets, design and sampling

Male C57BL/6J mice were purchased from Charles River Laboratories (Maastricht, the Netherlands) at 6 weeks of age. Mice were housed in pairs in a light- and temperature-controlled animal facility of Wageningen University (12 hour light-dark cycle; light on from 11h PM to 11h AM, 21 °C). Mice had free access to water and food throughout the entire experimental period. Upon arrival, mice were fed standard lab chow (RMH-B, Arie Blok, Woerden, the Netherlands) for 3 wks. Subsequently, all mice were adjusted to the control diet, a standard semi-synthetic low fat diet containing corn starch, for 2 wks. To achieve similar weight distribution among the diet groups, mice were stratified according to their body weight to one of the six diet groups (n=10 per diet group), i.e. control (CON), inulin (IN), oligofructose (FOS), arabinoxylan (AX), guar gum (GG) or resistant starch (RS). Mice were fed the fiber or control diets for 10 days. Detailed information on diet composition, procedure on the day of section and tissue sampling is described in **chapter 3**.

Short-chain fatty acid analysis in caecal luminal content

Short chain fatty acids were measured in mouse intestinal samples at section. Luminal content of the caecum (n=10 per group) was collected in H₃PO₄ and isocaproic acid (as an internal standard) containing buffer solution. Samples were stored at -20°C until further processing. The day of analysis, samples were thawed, centrifuged at 14.000 rpm (5 min), and supernatant was collected and stored at 5°C. The samples were then subjected to gas

chromatography (Fisons HRGC Mega 2, CE Instruments, Milan, Italy) at 190°C using a glass column fitted with Chromosorb 101 with a carrier gas (N₂ saturated with methanoic acid).

Microbial Composition

Total DNA was extracted from 0.01-0.1 grams of caecal content samples using the repeated bead beating plus column (RBB+C) method of [21]. The composition of microbial communities in the intestinal samples was analysed with the Mouse Intestinal Tract Chip (MITChip). This phylogenetic microarray was designed using criteria of the Human Intestinal Tract Chip (HITChip) [22]. The MITChip consists of 3,580 different oligonucleotides specific for the mouse intestinal microbiota [22], [23], [24]. The oligonucleotides on the array target the V1 and V6 regions of bacterial 16S rRNA genes. The 16S rRNA genes were amplified from twenty nanogram of DNA extracted from intestinal samples, with the primers T7prom-Bact-27-F and Uni-1492-R (**Table 1**). PCR products were then transcribed, and RNA was labelled with Cy3 and Cy5 dyes and fragmented as described previously [22], [23], [24]. Finally the samples were hybridized on the arrays at 62.5°C for 16h in a rotation oven (Agilent Technologies, Amstelveen, The Netherlands). After washing and scanning of the slides, data was extracted with the Agilent Feature Extraction software, version 9.1. The data was normalized and analysed using a set of R-based scripts in combination with a custom-designed relational database, which operates under the MySQL database management system. To determine correlation of the Robust Probabilistic Averaging (RPA) signal intensities of 2667 specific probes for the 96 genus-level bacterial groups detected on the MITChip with a specific diet or SCFA, redundancy analysis (RDA) in Canoco 5.0 was used, and visualized in a triplot [25], [26]. The Monte Carlo Permutation test was used to assess the significance of the variation in the dataset in relation to the diet and SCFA.

Table 1 List of primers (Suzuki et al. 2000, Rajilic-Stojanovic et al. 2009).

Primer name	Sequence	Application
T7prom-Bact-27-F	5'-TGA ATT GTA ATA CGA CTC ACT ATA GGG GTT TGA TCC TGG CTC AG-3'	MITChip
Uni-1492-R	5'-CGG CTA CCT TGT TAC GAC-3'	MITChip
PROK1492R	5'-GGW TAC CTT GTT ACG ACT T-3'	QPCR
BACT1369F	5'-CGG TGA ATA CGT TCY CGG-3'	QPCR

The Unpaired Wilcoxon signed-rank test was used to determine bacterial groups significantly different between the control (CON) and inulin (IN), oligofructose (FOS), arabinoxylan (AX), guar gum (GG) and resistant starch (RS) diet. The RPA signal intensities of the 96 genus-level groups were tested.

Quantification of bacterial community

Quantification of the bacterial 16S rRNA gene was done by a qPCR assay using the primers developed by [27]. The qPCRs were performed in 384-well plates (BioRad) sealed with a film (Microseal B film, Bio-Rad) using a MyIQ cycler with MyIQ software (version 1.0.410, Bio-Rad). The reactions were carried out in a total volume of 12.5 μ L consisting of 1x IQ SYBR green Supermix (Bio-Rad), 200 nM of the forward and reverse primer and 2 μ L of template DNA, and the cycling program and melting curve analysis as previously described [28]. The standard curve consisting of an 8-fold dilution series was a 16S rRNA gene PCR product of *Escherichia coli* top10. The copy number was calculated per caecum weight.

RNA extraction, mRNA enrichment, cDNA synthesis and Illumina sequencing

Four intestinal caecum content samples from each dietary treatment were used to analyse the metatranscriptome activity profiles. The RNA was extracted from 0.1-0.2 grams of caecal content. The content was re-suspended in 500 μ L ice-cold TE buffer (Tris-HCL pH 7.6, EDTA pH 8.0). Total RNA was obtained via the Macaloid-based RNA isolation protocol [29],[30] with in addition the use of Phase Lock Gel heavy tubes (5 Prime GmbH, Germany) during the phase separation. The RNA purification was done with the RNeasy mini kit (Qiagen, USA), including an on-column DNaseI (Roche, Germany) treatment [29]. The total RNA was eluted in 30 μ L ice-cold TE buffer and the RNA quantity and quality were assessed using a NanoDrop ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, USA) and Experion RNA Stdsens analysis kit (Biorad Laboratories Inc., USA), respectively. mRNA enrichment was performed by using the mRNA enrichment kit (MICROBExpress, Ambion, Applied Biosystem, the Netherlands) using the manufacturer's protocol. As with the total RNA the quantity and quality were assessed, the latter was done to check on the efficiency of the mRNA enrichment. One μ g of the enriched mRNA sample was used to transcribe the RNA in cDNA. Double stranded cDNA was synthesized with the SuperScript Double-Stranded cDNA Synthesis kit (Invitrogen, the Netherlands), with addition of SuperScript III Reverse Transcriptase (Invitrogen, the Netherlands) and random priming using random hexamers (Invitrogen, the Netherlands) [31],[32],[30]. To remove the RNA, a RNase A (Roche, Germany) treatment was preformed, followed by phenol-chloroform extraction of the cDNA and ethanol precipitation. The product was checked on gel and 3 to 8 μ g of cDNA was send to the sequencing provider for sequencing (GATC Biotech, Germany). Single read Illumina Libraries were prepared from the double-stranded cDNA according to the ChiP-seq protocol [33] with insert sizes between 200-300bp, suing barcoded tags for library constructions to enable parallel sequencing (GATC Biotech, Germany). Sequencing was performed using Illumina Hiseq2000 and using 5pM concentration of the library and the single-end protocol [30].

Sequence data processing

In total, sequencing yielded between 11 and 34 million reads per sample. The data is available in the NCBI small reads archive (sra) repository, under . The reads were processed via a previously described protocol (PhD thesis Floor Hugenholtz, Chapter 5). Briefly the data was filtered for ribosomal RNA sequences, adapter sequences and poor quality reads using SortMeRNA (version 1.2) [34], cutadapt [35], PRINSEQ (lite-version) [36], respectively. Reads smaller than 50 nucleotides were discarded. The resulting mRNA fractions were merged and de novo assembled into larger contigs, creating one reference set for all samples. A total of 70710 contigs were assembled with a total length of 59 Mb (n50=1074). Encoded in these contigs a total of 104110 potential open reading frames were predicted. To determine the taxonomic origin of the contigs, the predicted protein sequences were aligned with NCBI's non-redundant database, and the taxonomical family classification of the best hit was retrieved. Functional annotation was done by assigning KEGG orthology identifiers using the KEGG's KAAS server. Expression levels of the predicted proteins were determined by aligning the mRNA reads with assembled contigs and counting the total amount of nucleotides aligned with the corresponding ORFs.

Transcriptome analysis and Functional implications

All steps for Microarray hybridization and analysis, including RNA isolation and purification, were performed as described before [Chapter 3]. Functional implications were analysed using Enrichr [37].

Multivariate statistical analysis

We used Partial Least Square analysis (PLS) from mixOmics library in R to first, analyse effects of dietary fibers on gene expression (PLS Discriminant analysis) and second, to integrate microbiota composition and metatranscriptome with mucosal gene expression (PLS canonical correlation). For latter analysis, twenty-three animals were chosen for analysis of microbiota composition and host gene expression of which for 15 mice both datasets were available and 8 mice were added that were cage partner to increase the power. For transcriptome-transcriptome analysis 11 mice were included in the analysis of which for 7 mice both datasets were available and 4 mice were added that were cage partner (analysis was performed as described in chapter 3).

Results

The effects of different dietary fibers - resistant starch (RS), arabinoxylan (AX), fructooligosaccharides (FOS), inulin (IN) and guar gum (GG) - were studied to evaluate whether the responses of the microbiota and the host to different dietary fibers are congruent or diverse. The effects of these dietary interventions were measured in caecal samples by determination of the steady state levels of luminal SCFA, the microbiota composition by

16S rRNA gene profiling, and the microbiota activity by metatranscriptomics, respectively. In addition, the cognate host responses in the caecal mucosa were determined using murine microarrays and this was correlated to the 16S microbial composition dataset and to the microbial metatranscriptome data (see flow chart, **Figure 1**).

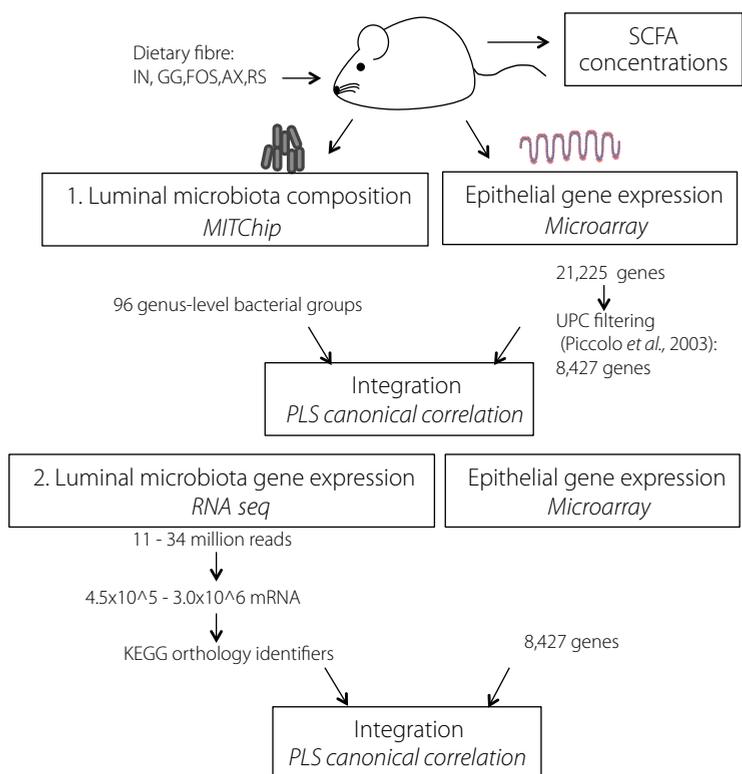


Figure 1 Flow-chart of data analysis

Dietary fibers differentially modulate luminal SCFA levels

In all mice the main metabolic products of dietary fiber fermentation in the caecal lumen were analysed by gas chromatographic quantification of acetate, propionate, butyrate, valerate, and the branched-chain SCFA iso-butyrate and iso-valerate. Total SCFA concentrations significantly increased in caecal lumen of mice in the IN and GG diet groups as compared to the CON diet ($p < 0.05$), and the highest SCFA concentrations were observed for mice fed the IN diet (**Figure 2**). In contrast, the RS, AX and FOS diets no changes the overall SCFA concentrations were observed.

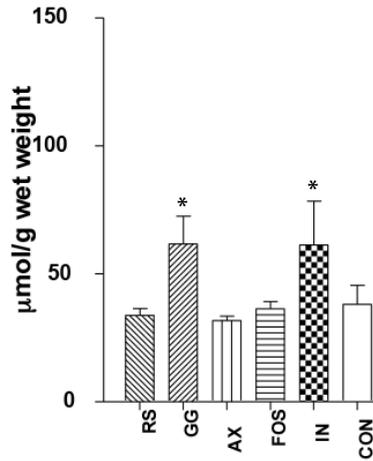


Figure 2 Ceacal luminal SCFA concentrations in $\mu\text{mol/g}$ content measured with gas chromatography

Abbreviations are for control (CON), Resistant Starch (RS), Arabinoxylan (AX), Fructooligosaccharides (FOS), Inulin (IN) and Guar Gum (GG). * indicates significance (Student T.test $P < 0.05$) between the dietary group and control (CON)

Dietary fibers modulate the microbiota composition

The caecal content of five mice per dietary treatment was used to analyse the microbiota after 10 days of dietary treatment. The density of the caecal microbiota was analysed by 16S rRNA gene-targeted quantitative PCR (qPCR), whereas the MITChip platform was employed for compositional profiling. Although all fiber diets, and especially the IN-diet, tended to increase the microbiota density in the caecum as compared to the CON diet, none of these effects were significant (**Figure S 1**). Microbial diversity was also not affected by the diets used in the present study (data not shown). Analogous to what has previously been reported for the colon in these mice, caecal microbiota composition of mice fed the GG, AX, IN, or FOS diet, but not those receiving an RS-diet, could be discriminated from that of mice receiving the control diet (**Figure S 2**). In order to correlate changes in microbiota composition and metabolism to the different diets, the SCFA concentrations and the hybridization signals of in total 96 genus-level phylogenetic groups targeted by the MITChip were subjected to redundancy analysis (RDA). The RDA included the concentrations of acetate, propionate and butyrate and the diets as explanatory variables, which were concluded to explain 58.6% of the total variation in microbial composition, of which 73.5% was captured within the first two canonical axes of the RDA analysis (**Figure 3**). The RS and CON diet groups clustered separately from the IN, AX, FOS and GG groups

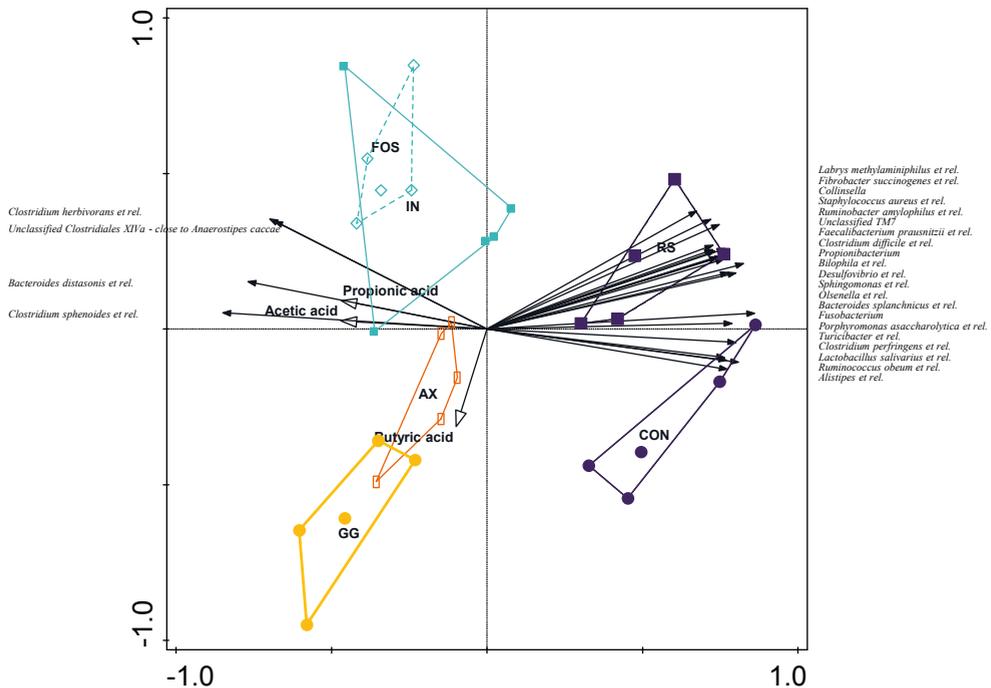


Figure 3

Redundancy analysis, RDA, where the explanatory variables are acetate, propionate, butyrate and the diets: control (CON), Resistant Starch (RS), Arabinosylin (AX), Fructooligosaccharides (FOS), Inulin (IN) and Guar Gum (GG). These variables explain 58.6% of total variation. In this plot 73.5% of the explained variation is shown.

along the first canonical axis, whereas the FOS and IN diet-groups appeared to be largely overlapping. Genus-like groups that correlated in their relative abundance with the RS and the control diet included *Collinsella*, *Propionibacterium*, *Olsenella* et rel., *Alistipes* et rel., *Bacteroides splanchnicus* et rel., *Porphyromonas asaccharolytica* et rel., *Fibrobacter succinogenes* et rel., *Lactobacillus salivarius* et rel., *Staphylococcus aureus* et rel., *Turicibacter* et rel., *Clostridium perfringens* et rel., *Faecalibacterium prausnitzii* et rel., *Ruminobacter amylophilus* et rel., *Clostridium difficile* et rel., *Ruminococcus obeum* et rel., *Fusobacterium*, *Bilophila* et rel., *Desulfovibrio* et rel., *Sphingomonas* et rel., *Labrys methylaminiphilus* et rel. and Unclassified TM7. In the opposite direction relative abundance of the groups *Clostridium herbivorans* et rel., *Clostridium sphenoides* et rel., an unclassified *Clostridium* cluster XIVa group and *Bacteroidetes distasonis* et rel. was correlated with IN, FOS, GG and AX diets. Moreover, luminal acetate and propionate concentration also correlated with the IN, FOS, GG and AX

diets and the four microbial groups that associated positively with these diets. An increased luminal butyrate concentration was observed with the GG and AX diets, but did not strongly correlate with a specific microbial group.

Effect of dietary fiber on caecum mucosal gene expression

We used Partial Least Square Discriminant Analysis (PLS DA) to analyse gene expression changes in caecal mucosa in mice fed different diets. The expression of 8427 filtered protein encoding genes was used as input for a sparse PLS DA. The analysis revealed that samples could mainly be distinguished between mice fed on CON and RS diets and mice fed IN, FOS and GG diets (**Figure 4**). AX fed mice were clustering between these two main clusters, indicating intermediate effects on gene expression as compared to all other diet groups. The data revealed that the different fiber diets modulate gene expression

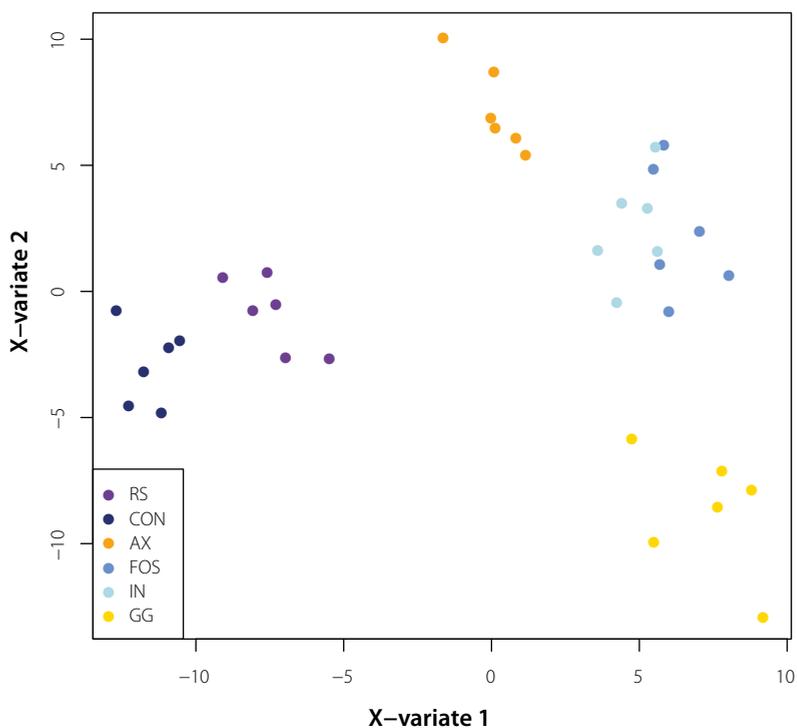


Figure 4

Partial Least Square Discriminant Analysis (PLS-DA) score plot on gene expression profiles in caecal epithelial cells of mice fed different diets: control (CON), resistant starch (RS), arabinoxylan (AX), fructooligosaccharides (FOS), inulin (IN) and guar gum (GG). In the plot the samples (individual mice) were plotted based on the two main variates.

differentially. This observation is similar to effects of dietary fiber on gene expression in colonic mucosa (**Chapter 3**), although contrary to the colon transcriptomes, the caecal tissue transcriptomes were clearly discriminated by diet, except for IN and FOS. This observation implies that fiber diets can elicit more specific effects on caecal mucosa as compared to colonic mucosa, which appears to be congruent with the notion that the caecum is the main fermentative organ in the mouse intestinal tract.

Correlation between microbiota composition and caecal mucosa gene expression profiles

To reveal associations between microbiota composition and mucosal gene expression profiles in the different diet groups, we integrated these two microarray datasets using a PLS-based canonical correlation approach. The results for the first three components were represented in a Clustered Image Map (CIM) with correlation coefficients depicted by different colours (**Figure 5**). In total 599 mucosal genes and 29 bacterial groups (abundance > 1%) were retained for the first three components, and clustering of the correlation coefficients revealed six main clusters of host genes that correlated to six main clusters of bacteria (Gene lists per cluster are available upon request). The largest microbiota cluster in this correlation analysis (**Figure 5**, cluster C) contained a range of bacterial groups, of which some were previously shown to be associated with the control and RS diets, like *Alistipes* et rel., *Bacteroides splanchnicus* et rel., *Porphyromonas asaccharolytica* et rel., *Clostridium difficile* et rel., *Ruminococcus obeum* et rel., *Desulfovibrio* et rel., *Sphingomonas* et rel., and Unclassified TM7. This microbiota cluster correlated with the repression of epithelial gene sets associated with phosphorylation, other signalling processes and protein modification (cluster 1), although this effect appeared not completely specific for this microbial cluster. Likewise, this microbial cluster also correlated with enhance expression levels of genes associated with protein catabolism and RNA related processes and apoptosis (cluster 4), although this host response was again not exclusively correlated to this group of bacteria. Stronger correlations were identified between the second largest bacterial cluster (cluster D) that contained microbial groups including *Parabacteroides distasonis* et rel., *Acitenobacter* et rel. and seven genus groups within the *Clostridium* cluster XIVa group. Remarkably, this microbiota cluster encompassed all the microbial groups that were previously found to be associated with the FOS, IN, AX, and GG diets, e.g. *Clostridium herbivorans* et rel., Unclassified *Clostridiales* XIVa (close to *Anaerostipes caccae*), *Clostridium sphenoides* et rel., and *Bacteroides distasonis* et rel. Cultured representatives of these bacterial groups are known for their capacity to degrade diet and host derived oligo- and poly-saccharides. Higher relative abundances of these bacterial groups was correlated with an elevated expression of genes associated with NF- κ B related processes, as well as proteolysis and energy-generating processes, including oxidative phosphorylation and glycolysis. Inversely, the relative abundance of these microbial groups correlated negatively with the expression of mucosal functions related to immune-

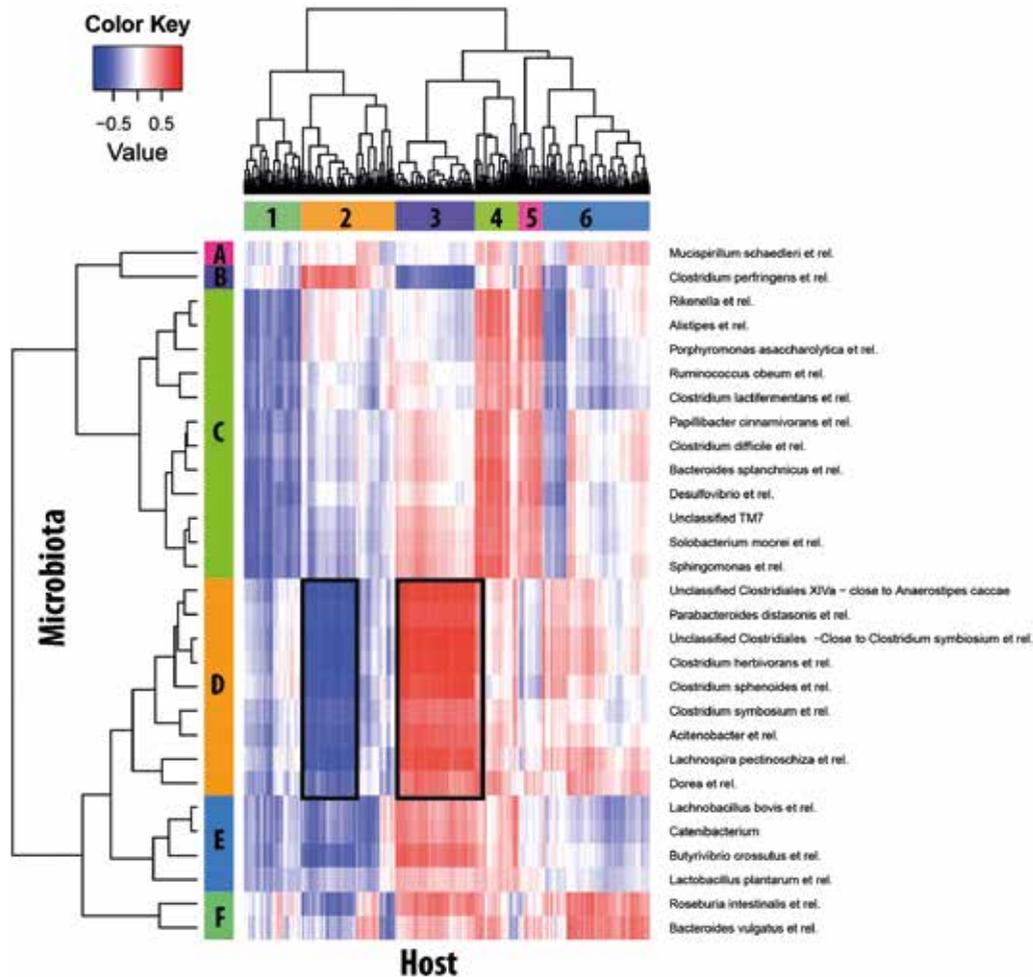


Figure 5

Correlation of epithelial cell gene expression with luminal microbiota composition. Sparse PLS canonical correlation analysis was performed to integrate gene expression values with relative abundance data of bacteria for individual mice. The heatmap represents the correlation structure of both dataset; red: positively correlated, blue: negatively correlated. More intense colours indicate stronger correlation. Correlation values were subjected to unsupervised hierarchical clustering based on Euclidean distance for both genes and microbial groups. Six main gene clusters (1-6) and six main bacterial clusters (A-F) were identified.

responses (T cell activation), DNA/RNA and glycosaminoglycan biosynthesis, and cell-cell communication. A single microbial group (*Clostridium perfringens* et rel.) appeared to correlate with exactly opposing mucosal transcriptome responses. The remaining microbial cluster A, E and F, encompassed functionally as well as phylogenetically broad microbial groups that correlated with relatively mild changes in caecal mucosa gene expression patterns.

These analyses illustrate that bacterial groups that were associated with the fiber containing diets (except the RS diet) and the elevated caecal SCFA concentrations, also were associated with alterations in gene expression profiles in the caecal mucosa of the mice in variable ways, affecting processes that include energy metabolism and immune response, which are the essential pillars of host-microbe homeostasis in the intestinal tract.

Effect of IN and GG on gene-functions expressed by the microbiota

The IN and GG diet groups showed the highest SCFA concentrations and most distinct microbial profiles compared to the CON diet. Therefore, the caecal content of 4 mice of the IN, GG and CON groups were used for metatranscriptome analysis, using a procedure encompassing extraction of total microbial RNA, rRNA depletion, double-strand cDNA synthesis and single-end shotgun Illumina sequencing. A total of 11 to 34 million metatranscriptome reads were generated per sample. These reads were filtered for non mRNA and low quality reads, leading to the extraction of 4.5×10^5 to 3.0×10^6 mRNA reads that were assembled and functionally and phylogenetically assigned to unravel the overall as well as group-specific microbiota activity profiles. Using this approach approximately 46-69% of the overall mRNA reads could be assigned to specific transcripts (**Table 2**), of which 70-85% was functionally annotated.

Using the sample specific functional assignments of the detected transcripts, without taking their phylogenetic origin into account, RDA analysis revealed that these function-patterns clustered according to diet, clearly separating CON, and GG and IN diets (**Figure 6**). To further unravel the functional impacts of the different fiber diets (GG and IN versus control), the functions (i.e., KEGGs) associated with the first axis of the RDA (explaining ~ 27.6 % of the overall variation within this analysis) were evaluated. Most of the differentially expressed KEGGs were within the KEGG category of "Metabolism" (>70%), predominated by carbohydrate, amino-acid, and central energy metabolism associated functions, as well as a scattering of other metabolic functions (**Table S 1**). In addition, genes belonging to the KEGG category "Environmental Information Processing" appeared to be substantially represented in the microbial activity patterns that discriminated the IN and GG diets from the CON diet. This category includes functions involved in response and adaptation of bacteria to their environment and encompasses both signal transduction pathways like two-component systems, as well as "membrane transport" associated functions like ABC-transporters and Phosphotransferase systems (PTS), that are

Table 2 Reads of the Illumina sequences and the result of data processing per sample.

	Total reads after quality filtering	mRNA	Assembled mRNA reads	Bacterial protein coding in assembled contigs
GG_1	1.3E+07	8.2E+05	61.3%	80.6%
CON_1	1.4E+07	9.4E+05	55.1%	83.6%
IN_1	1.7E+07	1.2E+06	54.6%	71.6%
GG_2	2.0E+07	1.8E+06	55.3%	84.9%
IN_2	1.6E+07	2.0E+06	58.6%	83.6%
GG_3	2.5E+07	1.7E+06	68.7%	75.8%
IN_3	1.4E+07	7.0E+05	58.4%	70.4%
CON_2	2.5E+07	1.5E+06	57.0%	79.4%
GG_4	2.0E+07	1.3E+06	45.7%	70.6%
CON_3	1.1E+07	4.5E+05	58.4%	79.8%
CON_4	2.8E+07	1.0E+06	54.0%	75.1%
IN_4	3.4E+07	3.0E+06	56.4%	83.0%

important for the utilization of alternative carbon sources (e.g., fibers from the diet), and are required to drive the adaptations of the overall microbiota metabolism that was detected. Intriguingly, the CON diet microbiota appeared to express genes with KEGG identifiers of the “cell motility” category at a higher level as compared to the microbiota of mice fed GG or IN diets. The induction of bacterial motility within the microbiota, specifically in the caeca of the CON-diet animals, may imply that the microbiota experiences nutrient starvation under these conditions, which appears in agreement with the lower abundance of dietary nutrients (e.g. fibers) in the caecal lumen of these animals, and is known to induce motility *in vitro* [38],[39].

Correlation between microbiota expressed functions and caecal mucosa gene expression patterns

To pinpoint microbial and host functions that are associated, the metatranscriptome KEGG functions were correlated with the caecal mucosa gene expression patterns using sPLS canonical correlations, leading to the detection of two strongly correlating co-clusters of bacterial and host transcripts. The strongest correlations (**Figure 7**, marked areas of the correlation analysis), encompassed host-functions associated with glutathione, glutamate metabolism and glycerolipid metabolic processes, and PPAR signalling, (negatively correlated; gene enrichment lists per cluster are available upon request), as well as immune-related processes such as B cell and T cell receptor and toll-like receptor signalling (positively correlated, gene enrichment lists per cluster are available upon request). The

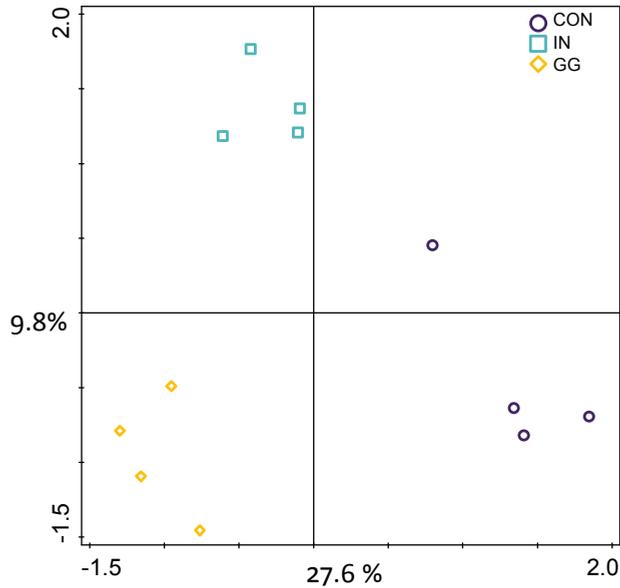


Figure 6

Redundancy analysis (RDA) of metatranscriptome data, where the explanatory variables are the diets control (CON), Inulin (IN) and Guar Gum (GG). These variables explain 37.4% of total variation. In this plot all the explained variation is shown.

microbial functions correlated with these host processes, were scattered across a variety of different microbial pathways and processes, and only displayed a minor enrichment of the KEGG category “Amino-acid metabolism”. Intriguingly, the microbiota-associated functions that displayed an opposite correlation with these host functions (**Figure 7**; the pink (B), purple (F) and dark blue (G) clusters), predominantly belonged to the KEGG categories “Environmental Information Processing”, “Cellular Processes”, “Genetic Information Processing” and “Nucleotide metabolism”.

Effect of IN and GG on active microbial community

The microbiota functional profiles determined in the different diet groups could be due to a shift in the relative contribution to the overall activity patterns by specific microbial groups rather than a function adaptation of the overall microbial community per se. The differentially expressed KEGG annotated functions could be assigned to five bacterial families, i.e., *Verrucomicrobiaceae*, *Bacteroidaceae*, *Lactobacillaceae*, and *Erysipelotrichaceae* were detected at a significantly higher level in the GG diet compared to the control diet (**Figure S3**). Conversely, the *Lachnospiraceae* assigned functions were in lower abundance

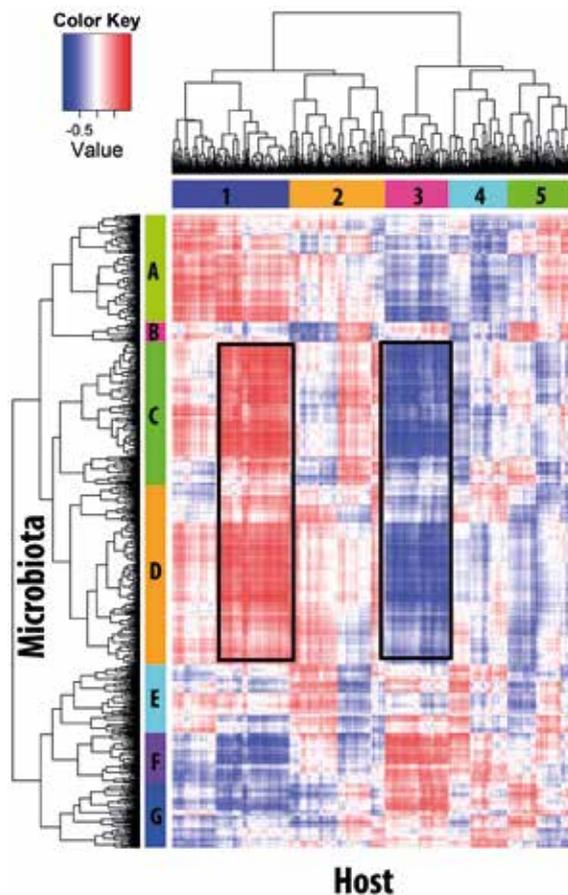


Figure 7

Correlation between microbiota expressed functions and caecal gene mucosa expression patterns. Sparse PLS canonical correlation analysis was performed to integrate epithelial cell gene expression with microbial metatranscriptome KEGG functions for individual mice. The heatmap represents the correlation structure of both datasets; red: positively correlated, blue: negatively correlated. More intense colours indicate stronger correlation. Correlation values were subjected to unsupervised hierarchical clustering based on Euclidean distance for both genes and microbial groups. Five main gene clusters (1-5), and seven main bacterial clusters (A-G) were identified.

in the GG diet compared to control. The same bacterial families displayed a similar trend in the IN samples compared to control samples, albeit not reaching significance. To identify which bacterial families are active in the degradation of the fibers, all metatranscriptome datasets were analysed in detail for the taxonomic origin of expressed

genes involved in glycoside hydrolysis (**Figure 8A**) and saccharide transport (**Figure 8BC**, **Table S 2**). Bifidobacteriaceae, Lachnospiraceae, Clostridiaceae and the Erysipelotrichaceae particularly expressed genes associated with glycoside hydrolysis and saccharide transport, possibly indicating active degradation of fibers and sugar transport into the cells. The *Bifidobacteriaceae* appeared to have increased expression of both glycosidases and sugar transporters in the IN and GG compared to the control. This family is known for its increase in abundance during interventions with oligofructosaccharides, like IN and FOS [40], [4], [18], [41]. The increased activity of *Bifidobacteriaceae* during GG intervention has not been reported to date, but according to the metatranscriptome data generated in this study appeared to be even somewhat more elevated in GG diet as compared to the IN diet. This finding implies that the *Bifidobacteriaceae* are strongly involved in GG fiber catabolism. The other bacterial families detected displayed quite distinct responses to the different diets. The *Lachnospiraceae* increased the glycosidase expression in IN and GG diets

A Glycosidases

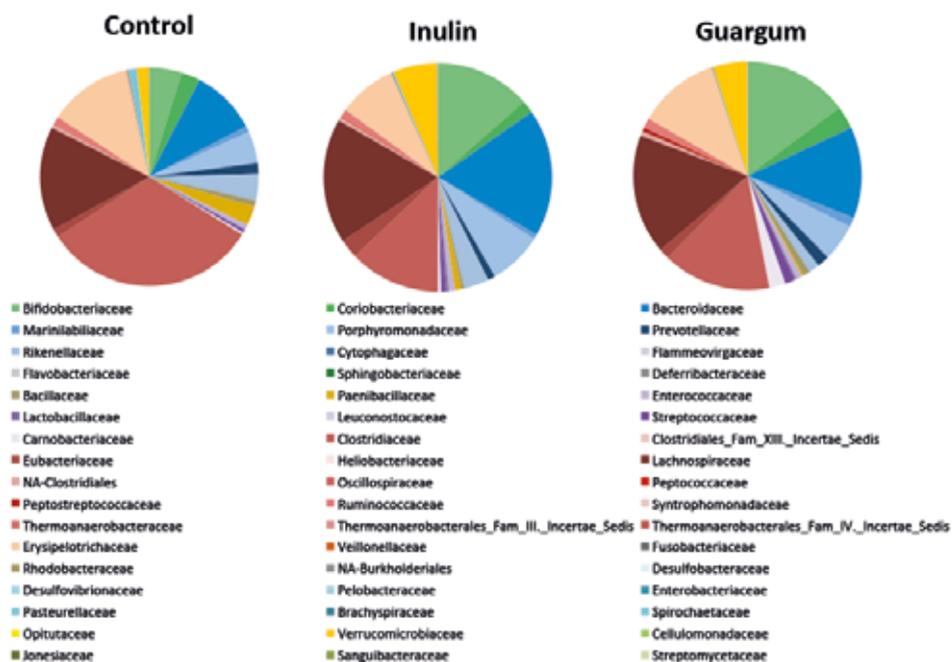


Figure 8 Relative abundance of gene expression of a) Glycosidases, b) ABC-transporters and c) Phosphotransferase systems (PTS) at family level.

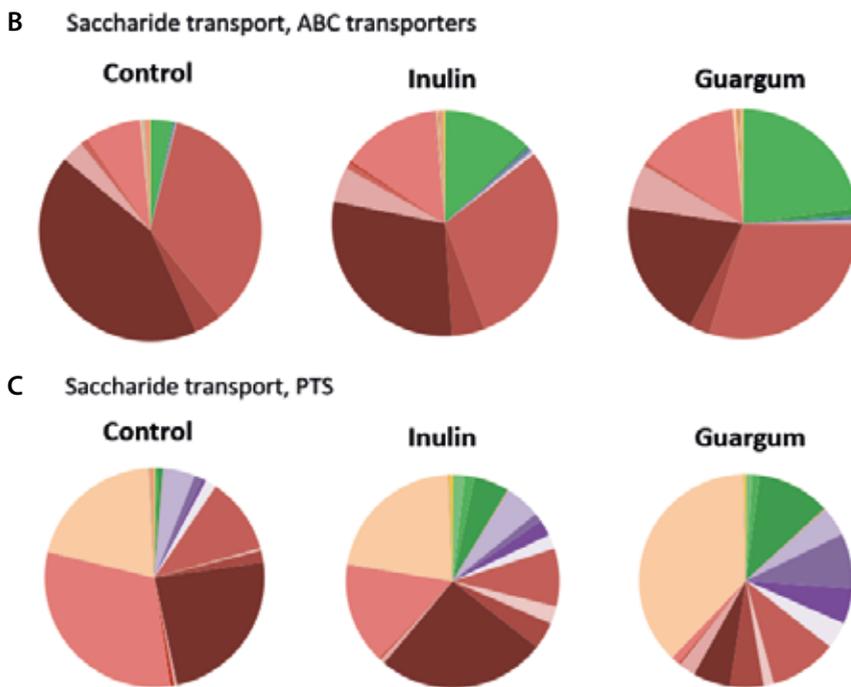


Figure 8 Continued.

compared to the control diet, while their expression of sugar transport functions was decreased. Conversely, the *Erysipelotrichaceae* elevated expression of sugar transport functions in the IN and GG diets, while their glycosidase expression appeared not to respond to the diet changes. Finally, the *Clostridiaceae* decreased the glycosidase expression levels in the fiber diets while the carbohydrate transport remained expressed at the same level in all diets.

Notably, although 25 % of all the detected glycosidase transcripts was assigned to the Bacteroidetes, this phylum hardly expressed saccharide transporter functions. The Bacteroidetes expressed a relatively large fraction of the overall (and diverse) hydrolase encoding genes, but appeared less active in importing the monomeric sugars produced by enzymatic polysaccharide hydrolysis. However, isolates within the *Bacteroides* and *Prevotella* genera encode operons that cluster genes encoding fiber binding, several carbohydrate degrading enzymes and transport functions [42], [43], and the encoded polysaccharide degradation machinery of *Bacteroides thetaiotaomicron* was proposed to be closely associated with the transport machinery anchored in the outer membrane. This assembly could ensure direct transfer from the hydrolytic enzymes to the transport

system to minimize loss of monosaccharides to other bacteria [42], which may not require elevated expression levels of other transport functions.

The Bacilli displayed a relatively lower expression of genes encoding glycosidases and higher expression of sugar transporter-encoding genes in the IN and GG groups. In particular the Phosphotransferase systems (PTS) expression by the *Bacilli* was strongly elevated in the IN and GG groups, suggesting that these populations depend on other microorganisms to generate the mono- and disaccharides through the initial hydrolysis of dietary polysaccharides. Analogously, also the *Ruminococcaceae* actively expressed sugar transport functions, but did not contribute to the expression of glycosidase encoding genes. This is somewhat unexpected, since members of the *Ruminococcaceae* are known as fiber degraders that possess glycosidases, like *Ruminococcus bromii*, *Faecalibacterium prausnitzii* and *Ruminococcus flavevacies* [44], [45]. Notably, the *Ruminococcaceae* shifted the expression of sugar transport systems from PTS systems to ABC transporters in the GG diet (and to a lesser extent in the IN diet) as compared to the CON diet. The substrate predictions of the PTS that are strongly expressed in the control diet group were related to saccharides that are derived from mucus degradation, which may be suppressed by the supply of dietary carbohydrates in the GG and IN diets (**Figure S 4**). PTS transport enables efficient import of mono- and/or disaccharides, which is coupled to substrate phosphorylation [44] and is commonly the preferred mode of transport in bacteria when they reside in carbohydrate limited environments. This is likely the situation for the *Ruminococcaceae* in the CON group caeca, where mucus-derived mono- and di-saccharides as well as other carbon sources were most likely depleted and/or not accessible for the members of this family. ABC transport allows import of mono- up to oligosaccharides, which could be energetically favourable for the *Ruminococcaceae* in the enriched environment associated with the fiber diets. This transport-flexibility is absent from many other bacterial families, like the *Lachnospiraceae*, of which the overall expression of carbon metabolism associated functions was suppressed in the GG and IN diet compared to the CON diet, implying that this bacterial family is unable to compete with other species for the dietary fibers.

With the intention to reconstruct the activity profiles involved in SCFA production, KEGGs assigned to key enzymes in SCFA production pathways were selected (**Figure S 5**). The expression of genes encoding these enzymes and their predicted taxonomic assignment differed between the CON, IN and GG groups in all of the pathways evaluated (**Figure 9**). In the GG diet the *Erysipelotrichaceae* seemed to more abundantly express various SCFA pathways compared to the control diet. The *Erysipelotrichaceae* were highly active in the conversion of pyruvate to lactate and/or vice-versa. In addition the *Erysipelotrichaceae* did express the acetyl-coA to butyryl-coA via crotonyl-coA pathway, where the conversion of crotonyl-coA to butyryl-coA allows anaerobes to conserve energy (**Figure S 6**; [46] and should lead to butyrate production. However the butyrate kinase and the butyryl-CoA:acetate CoA-transferase enzymes of *Erysipelotrichaceae* were not identified in our dataset,

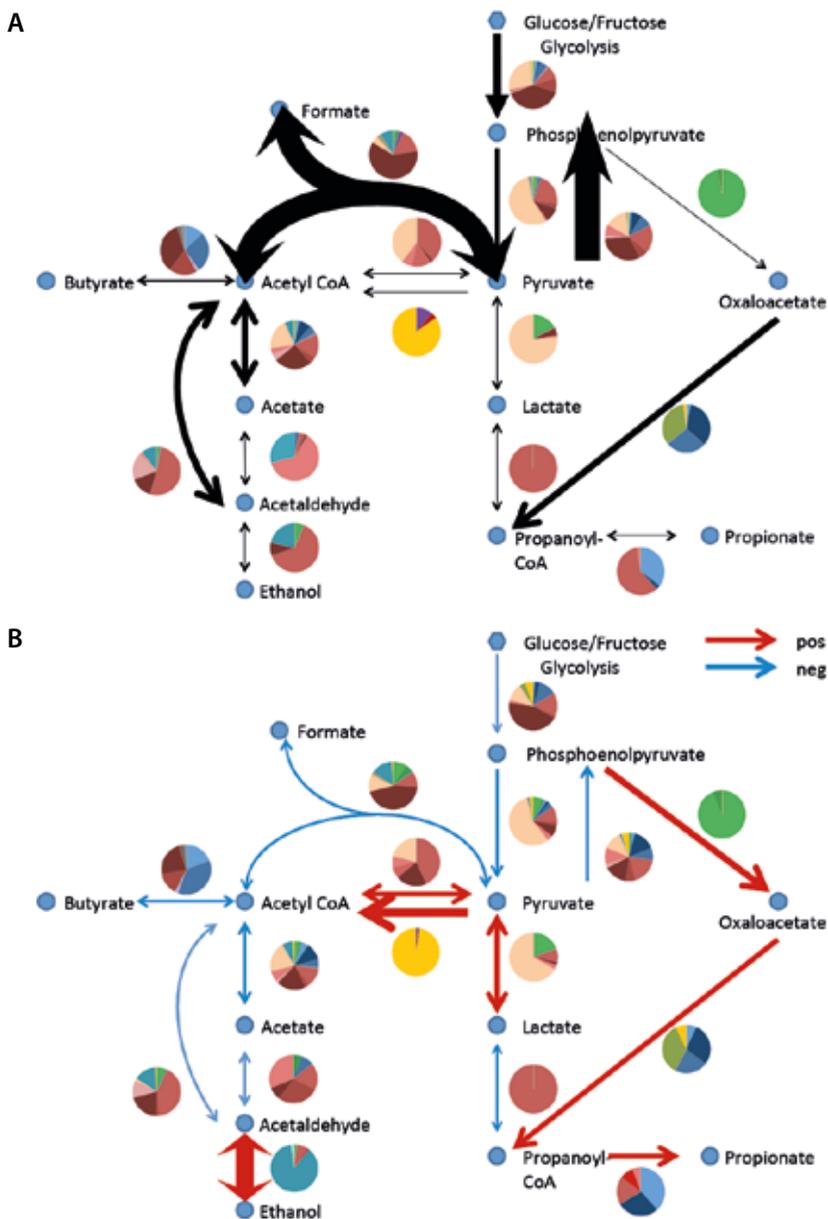


Figure 9 Activity of SCFA metabolism pathways

A) Control animals. Arrow thickness indicates activity of this pathway relative to the total activity. B) and C) are the Inulin and Guar gum dietary treatment, respectively. The colour and thickness of the arrows indicate a fold change compared to the same path in the control diet.

C

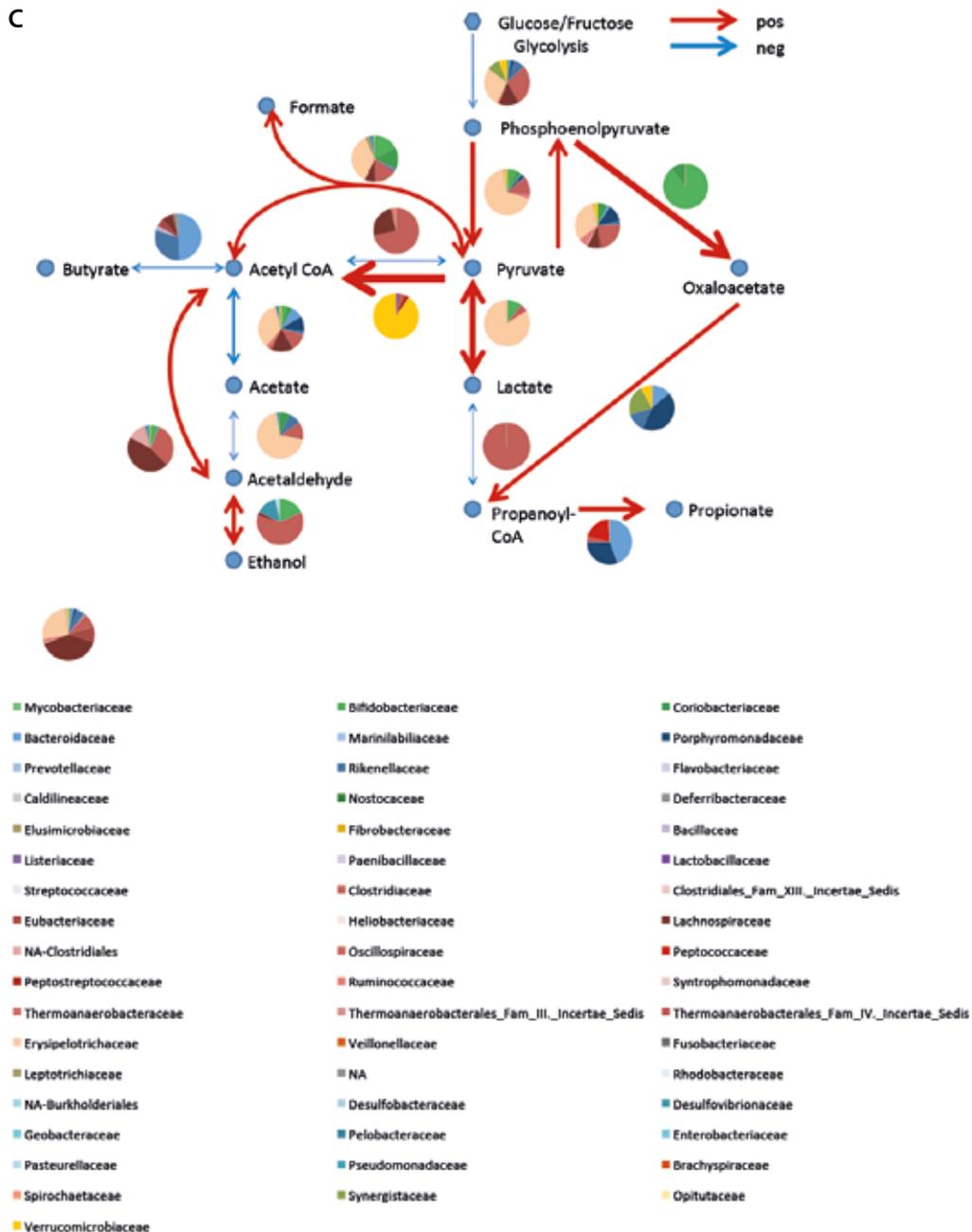


Figure 9 Continued

which may be due to the erroneous annotation of acetate and butyrate kinases and SCFA transferases since these enzymes display a high-level of similarity [47], [48,49]. Moreover, some of the phosphate butyryl-transferase and butyrate kinase were predicted to be expressed by members of the *Bacillaceae*, *Bacteroidaceae* and *Porphyromonadaceae* families, implying that these families are involved in the production of butyrate, which is in clear contrast with previous studies that indicated that *Bacillaceae* and *Bacteroidaceae* family members are not producing butyrate [45]. Again this may indicate that the inaccurate annotations of gene sequences related to kinases and transferases involved in SCFA pathways confuse these metabolic interpretations of the metatranscriptome data. In the IN group there was a high activity in ethanol consumption by members of the *Desulfovibrionaceae*, in line with previous reports that isolates within this family use ethanol as a carbon source [50]. Transcripts associated with the production of propionate were predominantly assigned to members of the *Clostridiaceae*, *Bacteroidaceae* and *Porphyromonadaceae*, where the latter two groups probably produce propionate from oxaloacetate via succinate. Although increased luminal concentrations of acetate, propionate and butyrate were detected in the GG and IN diets, only increased relative expression levels of genes associated with propionate were detected in the IN and GG metatranscriptomes compared to those of the CON diet, whereas transcripts associated with acetate and butyrate production appeared to be present at a relatively lower level in GG and IN compared to control diet.

Discussion

Dietary fibers have been associated with health benefits. After their degradation by the intestinal tract microbiota, SCFA are generated and taken up into the epithelium of the host. Here we show that the dietary fibers IN and GG yield increased total SCFA concentrations in the caecum and in parallel increase the expression of metabolic processes involved in central energy metabolism in the caecal mucosa of the mice. Moreover microbial analysis revealed shifts in composition and activity patterns when the mice were given the different carbohydrate sources.

Mice fed the FOS, AX or RS diets did not have increased total SCFA concentrations in their caeca, although FOS and AX could still stimulate enhanced expression of the central energy metabolism pathways in the mucosa of these mice. Similar observations were reported previously for the colonic mucosa of these mice, although in the colon lumen both the FOS and AX diets increased SCFA levels [chapter 3]. The experimental design of this study allowed sampling of luminal content at a single time point and thus generated a snapshot view of the intestinal system, which excludes interpretations of the rate of SCFA production and or consumption, which may be drastically affected by the different diets. Enhanced SCFA flux into the host mucosa could explain the similarity in the local mucosal responses measured in the IN, GG, FOS and AX diets. Such a snapshot

determination of SCFA concentrations in the lumen is clearly a measurement of limited value for the determination of the microbiota fibers-fermentation output [51]. Microbiota composition and activity shifts may provide a better proxy for the estimation of in situ fiber-fermentation rates in the intestinal lumen, and these parameters were found strongly affected by the addition of fibers to the diet.

The fiber digestion by the microbiota could be further specified by analysis of the expression of glycosidase and sugar transport functions. Here we could distinguish three categories of bacteria, (i) bacteria that express glycosidases, but hardly sugar transporters; (ii) bacteria that express both glycosidases and sugar transporters; (iii) bacteria that hardly express glycosidases, but do express sugar transporters (**Figure 10**). The first group is mainly represented by the *Bacteroidaceae*, the *Porphyromonadaceae* and the *Verrucomicrobioaceae*, of which the latter displayed elevated expression in the IN and GG diet. The *Verrucomicrobioaceae* member detected predominantly belonged to *Akkermansia muciniphila* (data not shown). The increase of activity in the IN and GG groups of this typical mucin digesting microbe [52] may indicate that these diets lead to increased mucus production, although this was not apparent from the mucosal transcriptome analyses, but may not be primarily regulated at the level of gene transcription [53]. The second group corresponds to members of the *Bifidobacteriaceae*, *Lachnospiraceae*, *Erysipelotrichaceae* and *Clostridiaceae*.

These four families all expressed sugar transporter encoding genes parallel to the glycosidases, which is in clear contrast to the first group and may be due to a lack of mechanistic coupling of polysaccharide hydrolysis and saccharide transport as has been proposed for some members of the first group of bacteria (e.g. *Bacteroides thetaiotamicron*; [42]). Notably, the *Erysipelotrichaceae* have only recently been recognized as a separate bacterial family, and many members still need to be characterized and re-assigned to this family [45], NCBI taxonomy, September 2014). Our results show that this family plays an important role in the murine microbiota and contributes strongly to its overall metabolite conversions (**Figure 10**). The third group are bacteria that profit from glycosidase activity of other bacteria and import the released sugars. These are the *Eubacteriaceae*, several Bacilli families and the *Ruminococcaceae*. All bacterial families in three different categories ferment the carbohydrates they ingest to produce SCFA in different composition and ratio, although their individual contribution to the overall SCFA production by the microbiota may differ as a function of the dietary treatment. However, deciphering of the specific activity in particular SCFA pathways of individual bacterial families was hampered by the apparent inaccuracy of SCFA pathway mapping of genes, where many functions appear to be wrongly assigned, due to the high degree of similarity of the enzymes involved. To overcome this, advanced annotation and domain recognition tools need to be developed to accurately dissect these different enzyme families, which is a prerequisite to enable SCFA pathway reconstruction for environmental samples on basis of metatranscriptome or similar metagenomic information.

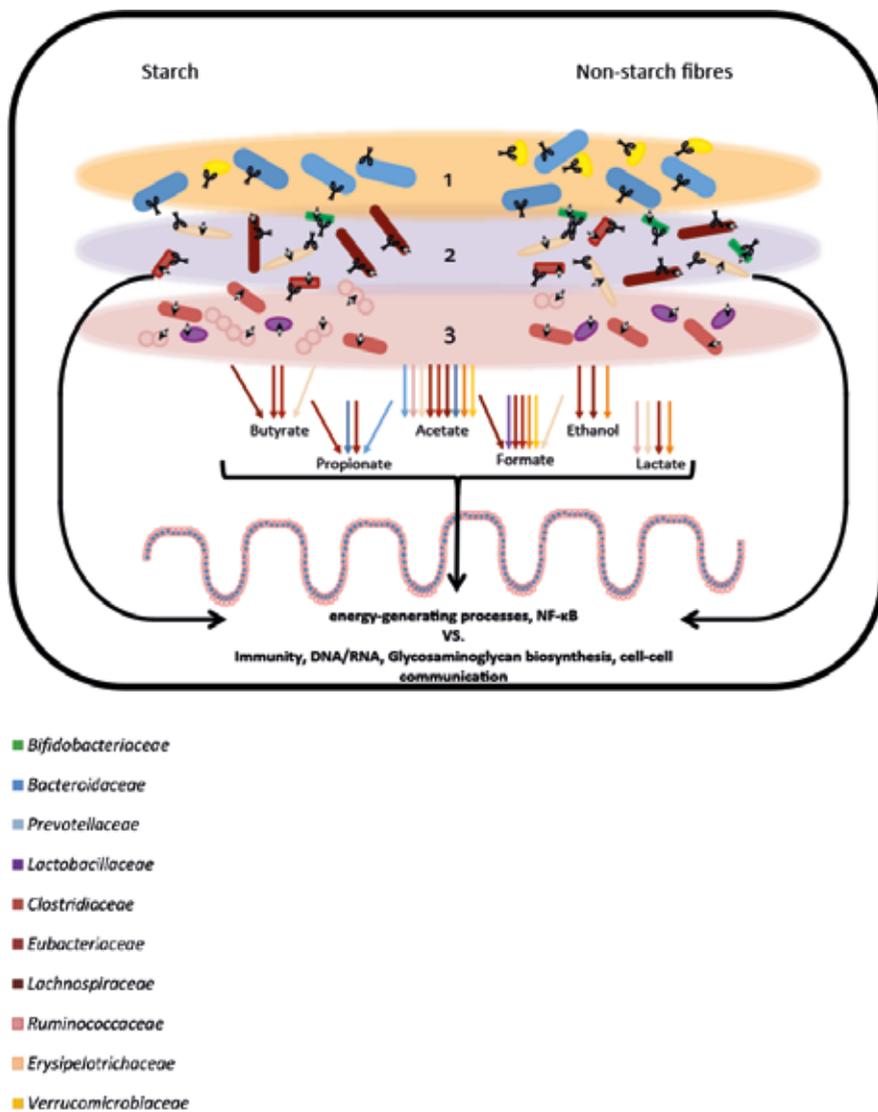


Figure 10 Overview of identified processes during fermentation

In summary the SCFA can be used as an energy source for the epithelial cells of the host. The correlation of MITChip and host gene-expression revealed fiber degrading and possibly butyrate producing bacteria activating energy metabolism in the host and repress transcriptional regulation and immune system processes. Moreover, based on the KEGG functions derived from the metatranscriptome data we observed a correlation with similar host genes to KEGG functions related to bacterial growth. This could indicate that active and fiber utilizing bacteria influence the host mucosa directly by enhancing its energy metabolism and affecting the immune system. Next to the known fiber responding families - *Bacteroidaceae*, *Porphyromonadaceae*, *Verrucomicrobioaceae*, *Bifidobacteriaceae*, *Lachnospiraceae*, *Clostridiaceae*, *Eubacteriaceae*, several *Bacilli* families and the *Ruminococcaceae* – we identified a new family, the *Erysipelotrichaceae*, as a prominent and active member of the murine gut microbiota.

Acknowledgments

We would like to thank Dr Diederick Meyer (Sensus) and Dr Hans van der Saag (BioActor) for their kind gifts of inulin and oligofructose, respectively arabinoxylan.

References

1. Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, et al. (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464: 59-65.
2. Tasse L, Bercovici J, Pizzut-Serin S, Robe P, Tap J, et al. (2010) Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes. *Genome research* 20: 1605-1612.
3. Flint HJ, Scott KP, Louis P, Duncan SH (2012) The role of the gut microbiota in nutrition and health. *Nat Rev Gastroenterol Hepatol* 9: 577-589.
4. Cummings JH, Macfarlane GT, Englyst HN (2001) Prebiotic digestion and fermentation. *The American journal of clinical nutrition* 73: 415S-420S.
5. Flint HJ, Bayer EA, Rincon MT, Lamed R, White BA (2008) Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nat Rev Microbiol* 6: 121-131.
6. Bird AR, Conlon MA, Christophersen CT, Topping DL (2010) Resistant starch, large bowel fermentation and a broader perspective of prebiotics and probiotics. *Beneficial microbes* 1: 423-431.
7. Flint HJ, Scott KP, Duncan SH, Louis P, Forano E (2012) Microbial degradation of complex carbohydrates in the gut. *Gut Microbes* 3: 289-306.
8. Mortensen PB, Clausen MR (1996) Short-chain fatty acids in the human colon: relation to gastrointestinal health and disease. *Scand J Gastroenterol Suppl* 216: 132-148.
9. Ouwehand AC, Derrien M, de Vos W, Tiihonen K, Rautonen N (2005) Prebiotics and other microbial substrates for gut functionality. *Current opinion in biotechnology* 16: 212-217.
10. Macfarlane S, Macfarlane GT, Cummings JH (2006) Review article: prebiotics in the gastrointestinal tract. *Alimentary pharmacology & therapeutics* 24: 701-714.
11. Al-Lahham S, Roelofsens H, Rezaee F, Weening D, Hoek A, et al. (2012) Propionic acid affects immune status and metabolism in adipose tissue from overweight subjects. *Eur J Clin Invest* 42: 357-364.
12. Donohoe Dallas R, Garge N, Zhang X, Sun W, O'Connell Thomas M, et al. (2011) The Microbiome and Butyrate Regulate Energy Metabolism and Autophagy in the Mammalian Colon. *Cell Metabolism* 13: 517-526.
13. Turnbaugh PJ, Gordon JI (2009) The core gut microbiome, energy balance and obesity. *J Physiol* 587: 4153-4158.
14. Cook SI, Sellin JH (1998) Review article: short chain fatty acids in health and disease. *Aliment Pharmacol Ther* 12: 499-507.
15. Bloemen JG, Venema K, van de Poll MC, Olde Damink SW, Buurman WA, et al. (2009) Short chain fatty acids exchange across the gut and liver in humans measured at surgery. *Clin Nutr* 28: 657-661.
16. Al-Lahham SH, Peppelenbosch MP, Roelofsens H, Vonk RJ, Venema K (2010) Biological effects of propionic acid in humans; metabolism, potential applications and underlying mechanisms. *Biochim Biophys Acta* 1801: 1175-1183.
17. Van den Abbeele P, Gerard P, Rabot S, Bruneau A, El Aidy S, et al. (2011) Arabinoxylans and inulin differentially modulate the mucosal and luminal gut microbiota and mucin-degradation in humanized rats. *Environ Microbiol* 13: 2667-2680.
18. Barboza M, Sela DA, Pirim C, Locascio RG, Freeman SL, et al. (2009) Glycoprofiling bifidobacterial consumption of galacto-oligosaccharides by mass spectrometry reveals strain-specific, preferential consumption of glycans. *Applied and environmental microbiology* 75: 7319-7325.
19. Riviere A, Moens F, Selak M, Maes D, Weckx S, et al. (2014) The ability of bifidobacteria to degrade arabinoxylan oligosaccharide constituents and derived oligosaccharides is strain dependent. *Applied and environmental microbiology* 80: 204-217.
20. Duncan SH, Louis P, Flint HJ (2004) Lactate-utilizing bacteria, isolated from human feces, that produce butyrate as a major fermentation product. *Applied and environmental microbiology* 70: 5810-5817.
21. Yu Z, Morrison M (2004) Improved extraction of PCR-quality community DNA from digesta and fecal samples. *Biotechniques* 36: 808-812.
22. Rajilic-Stojanovic M, Heilig HG, Molenaar D, Kajander K, Surakka A, et al. (2009) Development and application of the human intestinal tract chip, a phylogenetic microarray: analysis of universally conserved phylotypes in the abundant microbiota of young and elderly adults. *Environ Microbiol* 11: 1736-1751.
23. Geurts L, Lazarevic V, Derrien M, Everard A, Van Roye M, et al. (2011) Altered gut microbiota and endocannabinoid system tone in obese and diabetic leptin-resistant mice: impact on apelin regulation in adipose tissue. *Front Microbiol* 2: 149.

24. Reikvam DH, Derrien M, Islam R, Erofeev A, Grcic V, et al. (2012) Epithelial-microbial crosstalk in polymeric Ig receptor deficient mice. *Eur J Immunol* 42: 2959-2970.
25. Lahti L, Elo LL, Aittokallio T, Kaski S (2011) Probabilistic analysis of probe reliability in differential gene expression studies with short oligonucleotide arrays. *IEEE/ACM Trans Comput Biol Bioinform* 8: 217-225.
26. ter Braak CJF, P S (2012) Canoco reference manual and user's guide: software for ordination, version 5.0.: Microcomputer Power.
27. Suzuki MT, Taylor LT, DeLong EF (2000) Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5'-nuclease assays. *Applied and environmental microbiology* 66: 4605-4614.
28. van den Bogert B, de Vos WM, Zoetendal EG, Kleerebezem M (2011) Microarray analysis and barcoded pyrosequencing provide consistent microbial profiles depending on the source of human intestinal samples. *Applied and environmental microbiology* 77: 2071-2080.
29. Zoetendal EG, Booijink CC, Klaassens ES, Heilig HG, Kleerebezem M, et al. (2006) Isolation of RNA from bacterial samples of the human gastrointestinal tract. *Nature protocols* 1: 954-959.
30. Leimena MM, Ramiro-Garcia J, Davids M, van den Bogert B, Smidt H, et al. (2013) A comprehensive meta-transcriptome analysis pipeline and its validation using human small intestine microbiota datasets. *BMC genomics* 14: 530.
31. Yoder-Himes DR, Chain PS, Zhu Y, Wurtzel O, Rubin EM, et al. (2009) Mapping the Burkholderia cenocepacia niche response via high-throughput sequencing. *Proc Natl Acad Sci U S A* 106: 3976-3981.
32. Leimena MM, Wels M, Bongers RS, Smid EJ, Zoetendal EG, et al. (2012) Comparative analysis of *Lactobacillus plantarum* WCFS1 transcriptomes by using DNA microarray and next-generation sequencing technologies. *Applied and environmental microbiology* 78: 4141-4148.
33. Schmidt D, Wilson MD, Spyrou C, Brown GD, Hadfield J, et al. (2009) ChIP-seq: Using high-throughput sequencing to discover protein-DNA interactions. *Methods* 48: 240-248.
34. Kopylova E, Noé L, Touzet H (2012) SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* 28: 3211-3217.
35. Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet journal* 17: pp. 10-12.
36. Schmieder R, Edwards R (2011) Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27: 863-864.
37. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, et al. (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC bioinformatics* 14: 128.
38. Koirala S, Mears P, Sim M, Golding I, Chemla YR, et al. (2014) A Nutrient-Tunable Bistable Switch Controls Motility in *Salmonella enterica* Serovar Typhimurium. *MBio* 5: e01611-01614.
39. Aizawa SI, Kubori T (1998) Bacterial flagellation and cell division. *Genes to cells* 3: 625-634.
40. Gibson GR, Wang X (1994) Enrichment of bifidobacteria from human gut contents by oligofructose using continuous culture. *FEMS microbiology letters* 118: 121-127.
41. Kleerebezem M, Vaughan EE (2009) Probiotic and gut lactobacilli and bifidobacteria: molecular approaches to study diversity and activity. *Annual review of microbiology* 63: 269-290.
42. Dodd D, Mackie RI, Cann IK (2011) Xylan degradation, a metabolic property shared by rumen and human colonic Bacteroidetes. *Mol Microbiol* 79: 292-304.
43. Human Microbiome Jumpstart Reference Strains C, Nelson KE, Weinstock GM, Highlander SK, Worley KC, et al. (2010) A catalog of reference genomes from the human microbiome. *Science* 328: 994-999.
44. Biddle A, Stewart L, Blanchard J, Leschine S (2013) Untangling the Genetic Basis of Fibrolytic Specialization by Lachnospiraceae and Ruminococcaceae in Diverse Gut Communities. *Diversity* 5: 627-640.
45. Rajilic-Stojanovic M, de Vos WM (2014) The first 1000 cultured species of the human gastrointestinal microbiota. *FEMS Microbiol Rev.*
46. Seedorf H, Fricke WF, Veith B, Bruggemann H, Liesegang H, et al. (2008) The genome of *Clostridium kluverii*, a strict anaerobe with unique metabolic features. *Proc Natl Acad Sci U S A* 105: 2128-2133.
47. Louis P, Flint HJ (2007) Development of a semiquantitative degenerate real-time pcr-based assay for estimation of numbers of butyryl-coenzyme A (CoA) CoA transferase genes in complex bacterial samples. *Applied and environmental microbiology* 73: 2009-2012.

48. Vital M, Penton CR, Wang Q, Young VB, Antonopoulos DA, et al. (2013) A gene-targeted approach to investigate the intestinal butyrate-producing bacterial community. *Microbiome* 1: 8.
49. Vital M, Howe AC, Tiedje JM (2014) Revealing the bacterial butyrate synthesis pathways by analyzing (meta) genomic data. *MBio* 5: e00889.
50. Scanlan PD, Marchesi JR (2008) Micro-eukaryotic diversity of the human distal gut microbiota: qualitative assessment using culture-dependent and -independent analysis of faeces. *The ISME journal* 2: 1183-1193.
51. den Besten G, Havinga R, Bleecker A, Rao S, Gerding A, et al. (2014) The short-chain fatty acid uptake fluxes by mice on a guar gum supplemented diet associate with amelioration of major biomarkers of the metabolic syndrome. *PLoS One* 9: e107392.
52. Derrien M, Collado MC, Ben-Amor K, Salminen S, de Vos WM (2008) The Mucin degrader *Akkermansia muciniphila* is an abundant resident of the human intestinal tract. *Applied and environmental microbiology* 74: 1646-1648.
53. Hedemann MS, Theil PK, Bach Knudsen K (2009) The thickness of the intestinal mucous layer in the colon of rats fed various sources of non-digestible carbohydrates is positively correlated with the pool of SCFA but negatively correlated with the proportion of butyric acid in digesta. *British journal of nutrition* 102: 117-125.



Supplemental information

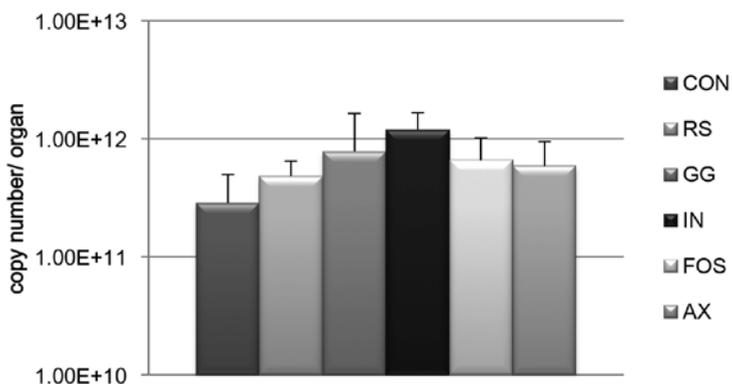


Figure S 1 Microbial abundance as measured by 16S rRNA gene-targeted quantitative PCR (qPCR)

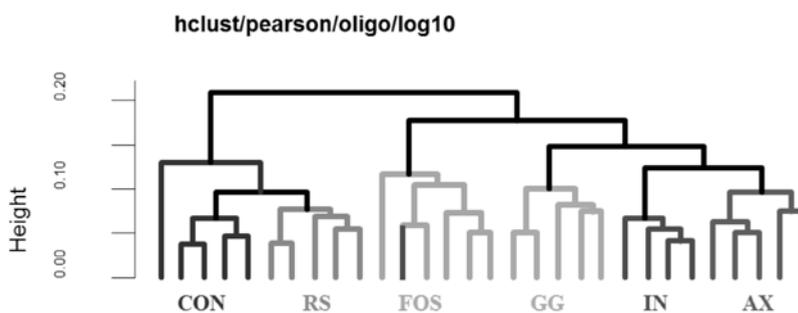


Figure S 2 Pearson distance clustering of the samples on log10 transformed probe level data of the MITChip

Abbreviations are for control (CON), Resistant Starch (RS), Arabinoxyylan (AX), Fructooligosaccharides (FOS), Inulin (IN) and Guar Gum (GG).

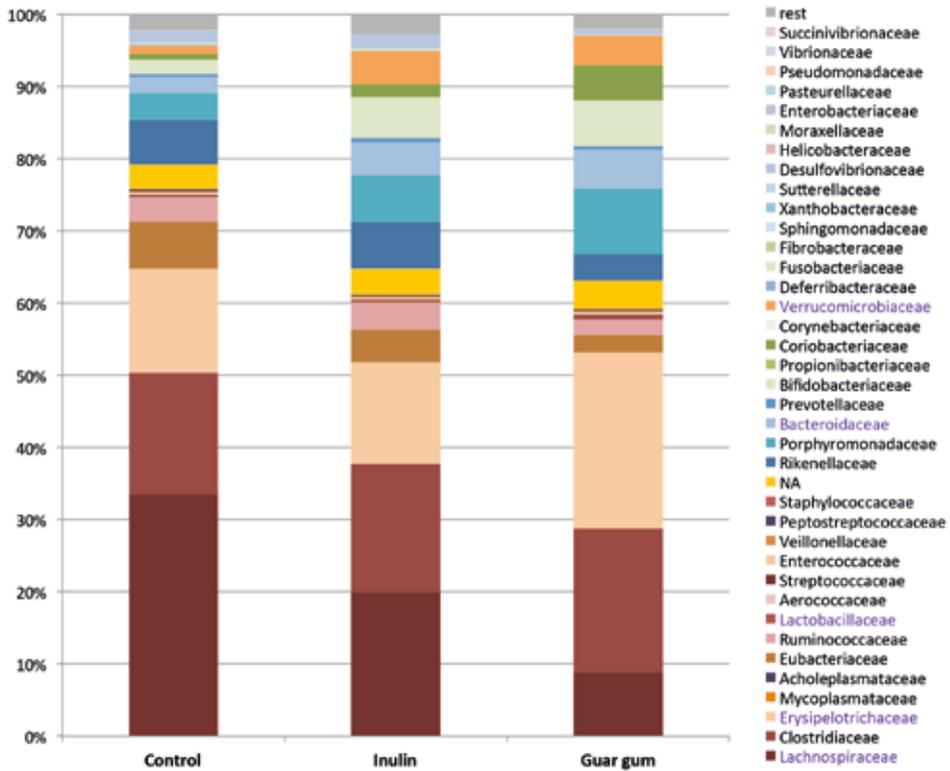
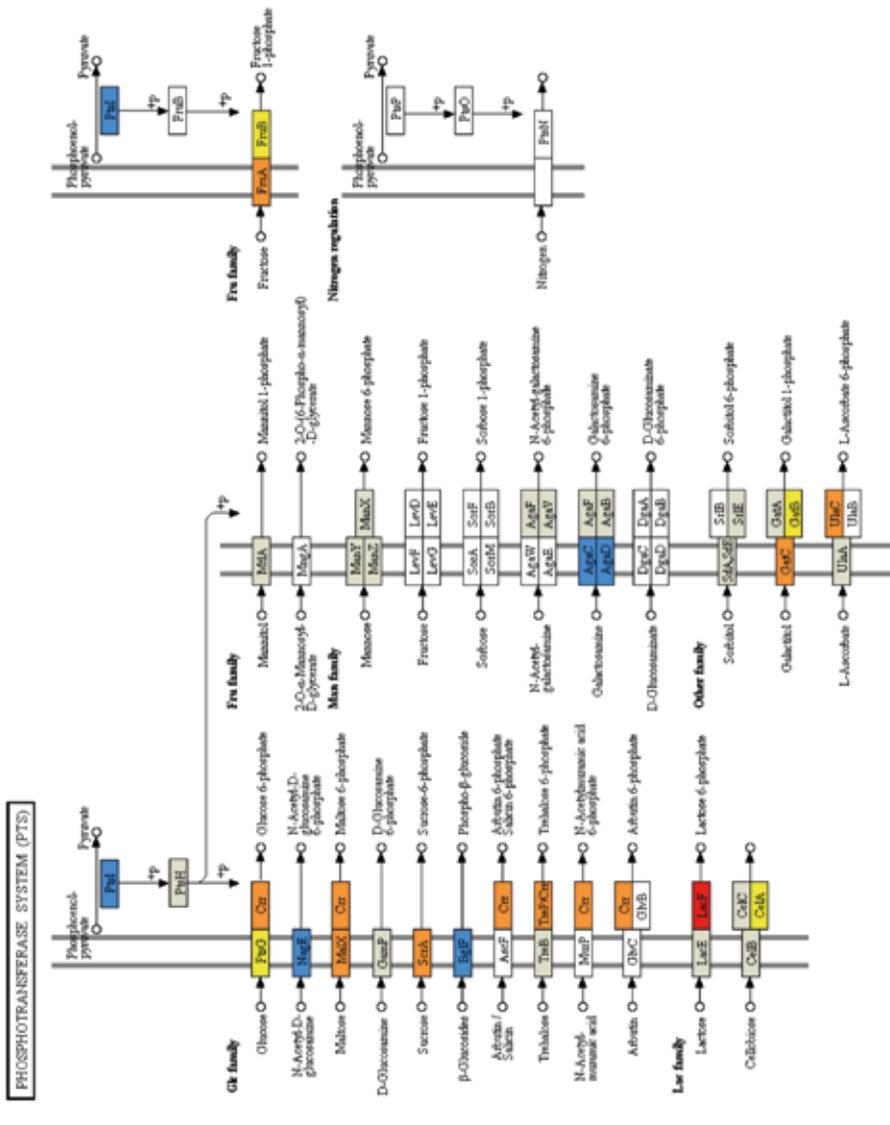


Figure S 3 Relative abundance of metatranscriptome (activity) at family level

Family names depicted in purple are significantly different in activity between control and Guar gum groups.



00060 01/013
(c) Kvarnåsen Laboratorier

Figure S 4 Expressed Phosphotransferase systems (PTS) in the metatranscriptome

Uncoloured units were not found in the data. Grey, equally expressed in all groups; orange, higher expression in IN and GG; blue, lower expression in IN and GG; yellow, higher expression in GG compared to CON; red, higher expression in IN compared to CON.

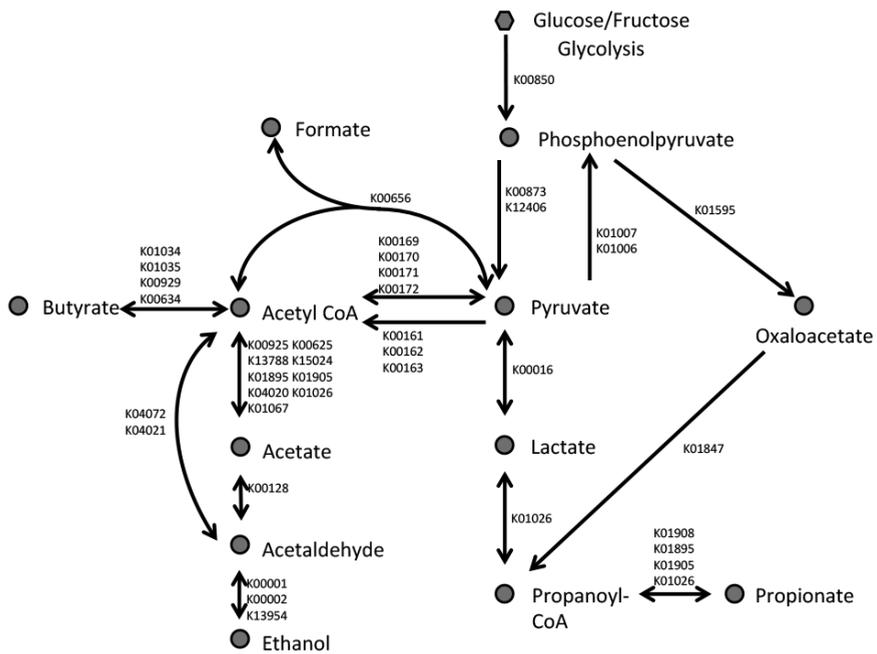


Figure S 5 KEGG numbers of the SCFA metabolism pathways used to create Figure 9

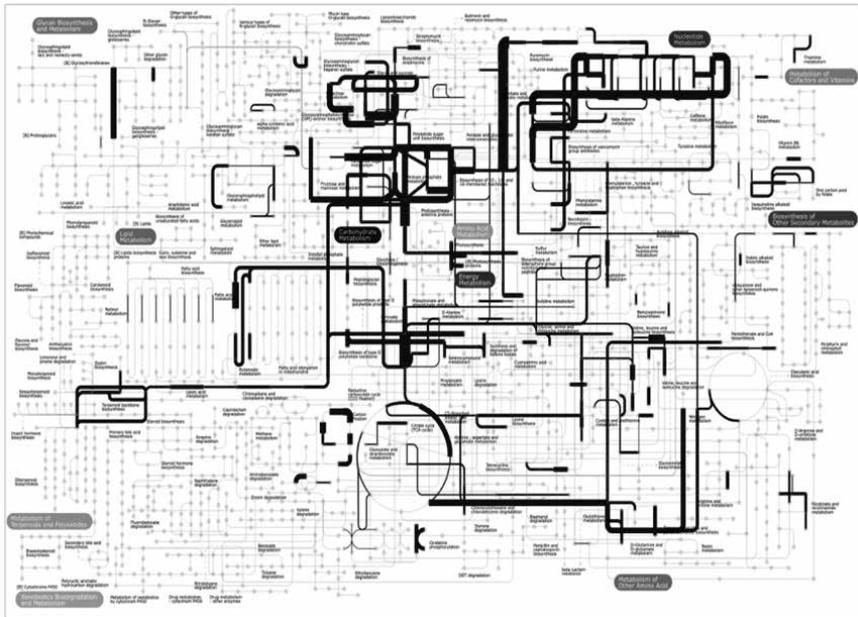


Figure S 6 iPATH visualization of the expression levels in the metabolic processes assigned to Erysipelotrichaceae

Table S 1 KEGG orthology distribution of 250 KEGGs explaining the differentiation of the diets in Figure 6.6 best

More expressed in IN and GG are the KEGGs in the direction of the IN and GG groups, less expressed are the KEGGs that correlated to the CON group.

	Nr of Keggs		Total Kegg
	<i>More expressed in IN and GG</i>	<i>Less expressed in IN and GG</i>	
Metabolism	95	87	70.5%
Carbohydrate metabolism	33.1%	38.6%	20.9%
Glycolysis / Gluconeogenesis	5	4	
Citrate cycle (TCA cycle)	4	-	
Pentose phosphate pathway	1	3	
Pentose and glucuronate interconversions	3	2	
Fructose and mannose metabolism	2	6	
Galactose metabolism	7	4	
Ascorbate and aldarate metabolism	1	1	
Starch and sucrose metabolism	2	-	
Amino sugar and nucleotide sugar metabolism	6	9	
Pyruvate metabolism	2	1	
Glyoxylate and dicarboxylate metabolism	2	1	
Propanoate metabolism	1	1	
Butanoate metabolism	4	2	
C5-Branched dibasic acid metabolism	1		
Energy metabolism	8.9%	6.7%	10.3%
Oxidative phosphorylation	3	-	
Carbon fixation in photosynthetic organisms	2	1	
Carbon fixation pathways in prokaryotes	-	1	
Methane metabolism	-	3	
Nitrogen metabolism	2	-	
Sulfur metabolism	4	3	
Lipid metabolism	6.5%	4.2%	3.8%
Fatty acid biosynthesis	-	1	
Fatty acid degradation	-	1	
Synthesis and degradation of ketone bodies	-	1	
Sphingolipid metabolism	2	1	
Linoleic acid metabolism	1	-	
Biosynthesis of unsaturated fatty acids	-	1	

Table S 1 Continued

	Nr of Keggs		Total Kegg
	<i>More expressed in IN and GG</i>	<i>Less expressed in IN and GG</i>	
Primary bile acid biosynthesis	1	-	
Glycerolipid metabolism	3	-	
Glycerophospholipid metabolism	1	-	
Nucleotide metabolism	6.5%	5.0%	7.4%
Purine metabolism	7	3	
Pyrimidine metabolism	1	3	
Amino acid metabolism	10.5%	7.6%	11.9%
Alanine, aspartate and glutamate metabolism	2	-	
Glycine, serine and threonine metabolism	1	1	
Cysteine and methionine metabolism	2	2	
Valine, leucine and isoleucine biosynthesis	1	-	
Valine, leucine and isoleucine degradation	-	1	
Lysine biosynthesis	1	-	
Lysine degradation	-	1	
Arginine and proline metabolism	4	1	
Phenylalanine, tyrosine and tryptophan biosynthesis	2	-	
Phenylalanine metabolism	-	2	
Tryptophan metabolism	-	1	
Metabolism of other amino acids	1.6%	1.7%	2.1%
Selenocompound metabolism	1	1	
D-Alanine metabolism	1	-	
beta-Alanine metabolism	-	1	
Glycan biosynthesis and metabolism	2.4%	5.0%	3.2%
N-Glycan biosynthesis	-	1	
Various types of N-glycan biosynthesis	-	2	
Glycosaminoglycan degradation	-	1	
Glycosphingolipid biosynthesis - globo series	1	-	
Peptidoglycan biosynthesis	1	-	
Other glycan degradation	1	2	
Metabolism of cofactors and vitamins	3.2%	5.9%	6.2%
Nicotinate and nicotinamide metabolism	1	-	
Porphyrin and chlorophyll metabolism	3	1	
Pantothenate and CoA biosynthesis	-	4	

Table S 1 Continued

	Nr of Keggs		Total Kegg
	<i>More expressed in IN and GG</i>	<i>Less expressed in IN and GG</i>	
Biotin metabolism	-	1	
Ubiquinone and other terpenoid-quinone biosynthesis	-	1	
Metabolism of terpenoids and polyketides	1.6%	0.8%	1.9%
Terpenoid backbone biosynthesis	1	1	
Carotenoid biosynthesis	1	-	
Biosynthesis of other secondary metabolites	0.8%	0.8%	1.2%
Streptomycin biosynthesis	1	-	
Caffeine metabolism	-	1	
Xenobiotics biodegradation and metabolism	1.6%	6.7%	1.6%
Chloroalkane and chloroalkene degradation	1	-	
Bisphenol degradation	1	-	
Benzoate degradation	-	3	
Xylene degradation	-	1	
Atrazine degradation	-	1	
Dioxin degradation	-	1	
Drug metabolism - other enzymes	-	2	
Genetic Information Processing	4	5	13.2%
Transcription	0%	0.8%	0.3%
RNA polymerase	-	1	
Translation	0.8%	2.5%	6.0%
Aminoacyl-tRNA biosynthesis	1	1	
Ribosome	-	2	
Folding, sorting and degradation	0.8%	0.8%	2.4%
RNA degradation	1	-	
Protein processing in endoplasmic reticulum	-	1	
Replication and repair	1.6%	0%	4.5%
Nucleotide excision repair	1	-	
Homologous recombination	1	-	
Environmental Information Processing	23	21	11.8%
Membrane transport	11.3%	11.8%	7.5%
ABC transporters	8	10	

Table S 1 Continued

	Nr of Keggs		Total Kegg
	<i>More expressed in IN and GG</i>	<i>Less expressed in IN and GG</i>	
Phosphotransferase system (PTS)	6	4	
Signal transduction	7.3%	5.9%	4.2%
Two-component system	6	6	
HIF-1 signaling pathway	1	-	
FoxO signaling pathway	1	-	
Phosphatidylinositol signaling system	1	-	
Calcium signaling pathway	-	1	
Cellular Processes	2	6	4.5%
Transport and catabolism	1.6%	0.8%	0.9%
Peroxisome	2	1	
Cell motility	0%	4.2%	2.9%
Bacterial chemotaxis	-	2	
Flagellar assembly	-	3	



Table S 2 Relative abundance of gene expression of Glycosidases, ABC-transporters and Phosphotransferase systems (PTS) at family level

A - glycosidases						
family	order	class	phylum	Control	Inulin	Guargum
Actinomycetaceae						
Bifidobacteriaceae	Bifidobacteriales{85004}	Actinobacteria{1760}	Actinobacteria{201174}	28657,3	108275,1	134188,9
Coriobacteriaceae	Coriobacteriales{84999}	Actinobacteria{1760}	Actinobacteria{201174}	16683,9	14206,4	29176,7
Bacteroidaceae	Bacteroidales{171549}	Bacteroidia{200643}	Bacteroidetes{976}	58608,5	144779,2	116977,6
Marinilibillaceae	Bacteroidales{171549}	Bacteroidia{200643}	Bacteroidetes{976}	5164,1	5390,1	11772,3
Porphyromonadaceae	Bacteroidales{171549}	Bacteroidia{200643}	Bacteroidetes{976}	28152,2	61078,5	47966,8
Prevotellaceae	Bacteroidales{171549}	Bacteroidia{200643}	Bacteroidetes{976}	9398,6	8186,9	15287,4
Rikenellaceae	Bacteroidales{171549}	Bacteroidia{200643}	Bacteroidetes{976}	23266,6	27850,3	13611,4
Flavobacteriaceae	Flavobacteriales{200644}	Flavobacteriia{117743}	Bacteroidetes{976}	2370	533,7	299,6
Nostocaceae				109,2	175,4	128,4
Ktedonobacteraceae	Ktedonobacterales{388448}	Ktedonobacteria{388447}	Chloroflexi{200795}	0,0	72,9	13,2
Deferribacteraceae	Deferribacterales{191393}	Deferribacteres{68337}	Deferribacteres{200930}			
Sphingobacteriaceae	Sphingobacteriales{200666}	Sphingobacteriia{117747}	Bacteroidetes{976}	61,7	160,2	75,9
Carnobacteriaceae	Lactobacillales{186826}	Bacilli{91061}	Firmicutes{1239}	1876,4	3601,2	20885,1
Bacillaceae	Bacillales{1385}	Bacilli{91061}	Firmicutes{1239}	5093,1	3027,8	8661,4
Enterococcaceae	Lactobacillales{186826}	Bacilli{91061}	Firmicutes{1239}	4350,7	7250,0	7656,9
Paenibacillaceae	Bacillales{1385}	Bacilli{91061}	Firmicutes{1239}	17331,7	7199,2	2436,1
Lactobacillaceae	Lactobacillales{186826}	Bacilli{91061}	Firmicutes{1239}	452,5	3095,2	1950,9
Streptococcaceae	Lactobacillales{186826}	Bacilli{91061}	Firmicutes{1239}	2776,0	5137,0	12902,1
Clostridiaceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	190821,3	103307,5	138745,4
Clostridiales						
Eubacteriaceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	7534,2	21643,2	13513,3
Hellobacteriaceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}			
Lachnospiraceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	91910,9	141298,7	152867,0

NA-Clostridiales	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	35070	4411,3	6101,7
Oscillospiraee	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}			
Peptococcaceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	481,1	645,9	1174
Peptostreptococcaceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	0,0	533,2	4655,1
Ruminococcaceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	6286,0	8080,7	127178
Halanaerobiaceae	Halanaerobiales{53433}	Clostridia{186801}	Firmicutes{1239}			
Thermoanaerobacteraceae	Thermoanaerobacteriales{68295}	Clostridia{186801}	Firmicutes{1239}	111,5	78,1	50,6
Thermoanaerobacterales_						
Fam_III_Incertae_Sedis				34,3	110,3	277
Thermoanaerobacterales_						
Fam_IV_Incertae_Sedis				308,4	223,2	0,0
Erysipelotrichaceae	Erysipelotrichales{526525}	Erysipelotrichia{526524}	Firmicutes{1239}	74378,5	65850,9	105011,4
Veillonellaceae	Selenomonadales{909929}	Negativicutes{909932}	Firmicutes{1239}	76,8	172,8	73,8
Fusobacteriaceae				771,9	24,6	0,0
Leptotrichiaceae						
NA						
Rhodobacteraceae	Rhodobacterales{204455}	Alphaproteobacteria{28211}	Proteobacteria{1224}	1171	286,6	208,5
Neisseriaceae	Neisseriales{206351}	Betaproteobacteria{28216}	Proteobacteria{1224}			
Desulfobacteraceae				140,0	30,0	68,5
Desulfovibrionaceae	Desulfvibrionales{213115}	Deltaproteobacteria{28221}	Proteobacteria{1224}	113,0	189,9	16,6
NA-Burkholderiales				75,2	115,6	71,4
Pelobacteraceae				0,0	93,7	68,5
Chromatiaceae				160,4	0,0	0,0
Pasteurellaceae				839,7	223,0	29,6
Pseudomonadaceae						
Brachyspiraceae	Spirochaetales{36}	Spirochaetia{203692}	Spirochaetes{203691}	339,4	995,6	910,0
Spirochaetaceae	Spirochaetales{36}	Spirochaetia{203692}	Spirochaetes{203691}	634,2	784,1	100,9



Table S 2 Continued

A - glycosidases						
family	order	class	phylum	Control	Inulin	Guargum
Synergistetes-NA						
Opiritaceae	NA	Opiritae{414999}	Verrucomicrobia{74201}			
Verrucomicrobiaceae	Verrucomicrobiales{48461}	Verrucomicrobiae{203494}	Verrucomicrobia{74201}	10719,7	49681,7	43530,8
Cellulomonadaceae	Actinomycetales{2037}	Actinobacteria{1760}	Actinobacteria{201174}	58,5	272,8	1070
Jonesiaceae	Actinomycetales{2037}	Actinobacteria{1760}	Actinobacteria{201174}	0,0	361,1	2074
Sanguibacteraceae	Actinomycetales{2037}	Actinobacteria{1760}	Actinobacteria{201174}	117,5	387,8	169,3
Streptomycetaceae	Actinomycetales{2037}	Actinobacteria{1760}	Actinobacteria{201174}	110,3	265,0	13,6
Leuconostocaceae	Lactobacillales{186826}	Bacilli{91061}	Firmicutes{1239}	0,0	74,1	166,8

B - ABC						
family	order	class	phylum	Control	Inulin	Guargum
Actinomycetaceae				266,6	157,7	576
Bifidobacteriaceae	Bifidobacteriales{85004}	Actinobacteria{1760}	Actinobacteria{201174}	39301,7	124289,9	187695,9
Coriobacteriaceae	Coriobacteriales{84999}	Actinobacteria{1760}	Actinobacteria{201174}	2454,5	5374,5	9329,5
Bacteroidaceae	Bacteroidales{171549}	Bacteroidia{200643}	Bacteroidetes{976}	227,3	412,2	1476,8
Marinilibiaceae	Bacteroidales{171549}	Bacteroidia{200643}	Bacteroidetes{976}			
Porphyrimonadaceae	Bacteroidales{171549}	Bacteroidia{200643}	Bacteroidetes{976}	385,6	642,2	1536,5
Prevotellaceae	Bacteroidales{171549}	Bacteroidia{200643}	Bacteroidetes{976}			
Rikenellaceae	Bacteroidales{171549}	Bacteroidia{200643}	Bacteroidetes{976}	2185,1	3140,4	1423,7
Flavobacteriaceae	Flavobacteriales{200644}	Flavobacteria{117743}	Bacteroidetes{976}			
Nostocaceae						
Ktedonobacteraceae	Ktedonobacterales{388448}	Ktedonobacteria{388447}	Chloroflexi{200795}	0,0	40,7	51,0

Deferribacteraceae	Deferribacterales{191393}	Deferribacteres{68337}	Deferribacteres{200930}		
Sphingobacteriaceae	Sphingobacteriales{200666}	Sphingobacteriia{117747}	Bacteroidetes{976}		
Carnobacteriaceae	Lactobacillales{186826}	Bacilli{91061}	Firmicutes{1239}	38,8	148,6
Bacillaceae	Bacillales{385}	Bacilli{91061}	Firmicutes{1239}	92,7	531,9
Enterococcaceae	Lactobacillales{186826}	Bacilli{91061}	Firmicutes{1239}	261,2	1732,5
Paenibacillaceae	Bacillales{385}	Bacilli{91061}	Firmicutes{1239}	1359,5	1595,2
Lactobacillaceae	Lactobacillales{186826}	Bacilli{91061}	Firmicutes{1239}	0,0	240,1
Streptococcaceae	Lactobacillales{186826}	Bacilli{91061}	Firmicutes{1239}	3289	2863,6
Clostridiaceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	433546,9	291263,3
Clostridiales				111,6	58,1
Eubacteriaceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	49080,3	45581,0
Hellobacteriaceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}		23704,1
Lachnospiraceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	523422,2	283975,8
NA-Clostridiales	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	39748,7	48575,2
Oscillospiraceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	10640,9	9091,7
Peptococcaceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	450	230,5
Peptostreptococcaceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	14399	232,1
Ruminococcaceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	100730,8	137921,8
Halanaerobiaceae	Halanaerobiales{53433}	Clostridia{186801}	Firmicutes{1239}	299	589
Thermoanaerobacteraceae	Thermoanaerobacterales{68295}	Clostridia{186801}	Firmicutes{1239}	583,1	2539,5
Thermoanaerobacterales_					
Fam_III_Incertae_Sedis				27,5	526,1
Thermoanaerobacterales_					
Fam_IV_Incertae_Sedis				34174	2402,0
Erysipelotrichaceae	Erysipelotrichales{526525}	Erysipelotrichia{526524}	Firmicutes{1239}	842,5	1400,1
Veillonellaceae	Selenomonadales{909929}	Negativicutes{909932}	Firmicutes{1239}		19299



Table S 2 Continued

B - ABC						
family	order	class	phylum	Control	Inulin	Guargum
Fusobacteriaceae						
Leptotrichiaceae						
NA						
Rhodobacteraceae	Rhodobacterales[204455]	Alphaproteobacteria[28211]	Proteobacteria[1224]	23.6	630.9	884.3
Neisseriaceae	Neisseriales[206351]	Betaproteobacteria[28216]	Proteobacteria[1224]	355.5	271	106.6
Desulfobacteraceae						
Desulfovibrionaceae	Desulfovibrionales[213115]	Deltaproteobacteria[28221]	Proteobacteria[1224]	327.0	624.5	201.5
NA-Burkholderiales			Proteobacteria[1224]			
Pelobacteraceae						
Chromatiaceae				1836.8	37.4	193.0
Pasteurellaceae						
Pseudomonadaceae						
Brachyspiraceae	Spirochaetales[136]	Spirochaetia[203692]	Spirochaetes[203691]	39.2	363.1	0.0
Spirochaetaceae	Spirochaetales[136]	Spirochaetia[203692]	Spirochaetes[203691]	10123.5	4786.8	2537.2
Synergistetes-NA				160.4	604.6	46.7
Opitutaceae	NA	Opitutae[414999]	Verrucomicrobia[74201]	119.8	79.5	166.1
Verrucomicrobiaceae	Verrucomicrobiales[48461]	Verrucomicrobiae[203494]	Verrucomicrobia[74201]	437.6	1193.1	919.5
Cellulomonadaceae	Actinomycetales[2037]	Actinobacteria[1760]	Actinobacteria[201174]			
Jonesiaceae	Actinomycetales[2037]	Actinobacteria[1760]	Actinobacteria[201174]			
Sanguibacteraceae	Actinomycetales[2037]	Actinobacteria[1760]	Actinobacteria[201174]			
Streptomycetaceae	Actinomycetales[2037]	Actinobacteria[1760]	Actinobacteria[201174]			
Leuconostocaceae	Lactobacillales[186826]	Bacilli[91061]	Firmicutes[1239]			

C - PTS

family	order	class	phylum	Control	Inulin	Guargum
Actinomycetaceae				530,3	9266,1	5982,7
Bifidobacteriaceae	Bifidobacteriales{85004}	Actinobacteria{1760}	Actinobacteria{201174}	3233,9	7401,2	6937,8
Coriobacteriaceae	Coriobacteriales{84999}	Actinobacteria{1760}	Actinobacteria{201174}	5515,4	26638,5	68934,9
Bacteroidaceae	Bacteroidales{171549}	Bacteroidia{200643}	Bacteroidetes{976}	15,1	349,5	304,4
Mariniliballiaceae	Bacteroidales{171549}	Bacteroidia{200643}	Bacteroidetes{976}			
Porphyromonadaceae	Bacteroidales{171549}	Bacteroidia{200643}	Bacteroidetes{976}	0,0	68,4	311,3
Prevotellaceae	Bacteroidales{171549}	Bacteroidia{200643}	Bacteroidetes{976}			
Rikenellaceae	Bacteroidales{171549}	Bacteroidia{200643}	Bacteroidetes{976}			
Flavobacteriaceae	Flavobacteriales{200644}	Flavobacteriia{117743}	Bacteroidetes{976}			
Nostocaceae						
Ktedonobacteraceae	Ktedonobacterales{388448}	Ktedonobacteria{388447}	Chloroflexi{200795}			
Deferribacteriaceae	Deferribacterales{191393}	Deferribacteres{68337}	Deferribacteres{200930}	159,2	68,1	64,7
Sphingobacteriaceae	Sphingobacteriales{200666}	Sphingobacteriia{117747}	Bacteroidetes{976}			
Carnobacteriaceae	Lactobacillales{186826}	Bacilli{91061}	Firmicutes{1239}	227,6	594,2	1082,3
Bacillaceae	Bacillales{1385}	Bacilli{91061}	Firmicutes{1239}	34518,9	26935,0	27750,7
Enterococcaceae	Lactobacillales{186826}	Bacilli{91061}	Firmicutes{1239}	8719,5	6427,6	52244,9
Paenibacillaceae	Bacillales{1385}	Bacilli{91061}	Firmicutes{1239}			
Lactobacillaceae	Lactobacillales{186826}	Bacilli{91061}	Firmicutes{1239}	4493,0	12283,8	33615,5
Streptococcaceae	Lactobacillales{186826}	Bacilli{91061}	Firmicutes{1239}	12037,0	10921,9	25258,6
Clostridiaceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	83325,7	44438,7	63656,5
Clostridiales				2334,1	13310,0	9660,8
Eubacteriaceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	13042,6	20600,8	31492,8
Hellobacteriaceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}			
Lachnospiraceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	176294,6	128565,4	35401,7
NA-Clostridiales	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	3817,3	4082,2	16118,0



Table S 2 Continued

C - PTS						
family	order	class	phylum	Control	Inulin	Guargum
Oscillospiraceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}			
Peptococcaceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}			
Peptostreptococcaceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	3048,9	902,7	1266,3
Ruminococcaceae	Clostridiales{186802}	Clostridia{186801}	Firmicutes{1239}	228942,0	77445,3	8323,6
Halanaerobiaceae	Halanaerobiales{53433}	Clostridia{186801}	Firmicutes{1239}			
Thermoanaerobacteraceae	Thermoanaerobacterales{68295}	Clostridia{186801}	Firmicutes{1239}	993,6	676,7	0,0
Thermoanaerobacterales_						
Fam_III_Incertae_Sedis						
Thermoanaerobacterales_						
Fam_IV_Incertae_Sedis						
Erysipelotrichaceae	Erysipelotrichales{526525}	Erysipelotrichia{526524}	Firmicutes{1239}	150426,9	109702,3	234281,4
Veillonellaceae	Selenomonadales{909929}	Negativicutes{909932}	Firmicutes{1239}	2779	505,9	360,5
Fusobacteriaceae						
Leptotrichiaceae						
NA						
Rhodobacteraceae	Rhodobacterales{204455}	Alphaproteobacteria{28211}	Proteobacteria{1224}			
Neisseriaceae	Neisseriales{206351}	Betaproteobacteria{28216}	Proteobacteria{1224}			
Desulfobacteraceae						
Desulfovibrionaceae	Desulfvibrionales{213115}	Deltaproteobacteria{28221}	Proteobacteria{1224}			
NA-Burkholderiales						
Pelobacteraceae						
Chromatiaceae				411,4	465,6	11,4
Pasteurellaceae				0,0	23,5	209,4
Pseudomonadaceae						

Brachyspiraceae	Spirochaetales[136]	Spirochaetia[203692]	Spirochaetes[203691]	4801,8	0,0
Spirochaetaceae	Spirochaetales[136]	Spirochaetia[203692]	Spirochaetes[203691]	4801,8	0,0
Synergistetes-NA					
Opitutaceae	NA	Opitutae[414999]	Verrucomicrobia[74201]		
Verrucomicrobiaceae	Verrucomicrobiales[48461]	Verrucomicrobiae[203494]	Verrucomicrobia[74201]	518,1	2263,1
Cellulomonadaceae	Actinomycetales[2037]	Actinobacteria[1760]	Actinobacteria[201174]		171,1
Jonesiaceae	Actinomycetales[2037]	Actinobacteria[1760]	Actinobacteria[201174]		
Sanguibacteraceae	Actinomycetales[2037]	Actinobacteria[1760]	Actinobacteria[201174]		
Streptomycetaceae	Actinomycetales[2037]	Actinobacteria[1760]	Actinobacteria[201174]		
Leuconostocaceae	Lactobacillales[186826]	Bacilli[91061]	Firmicutes[1239]		





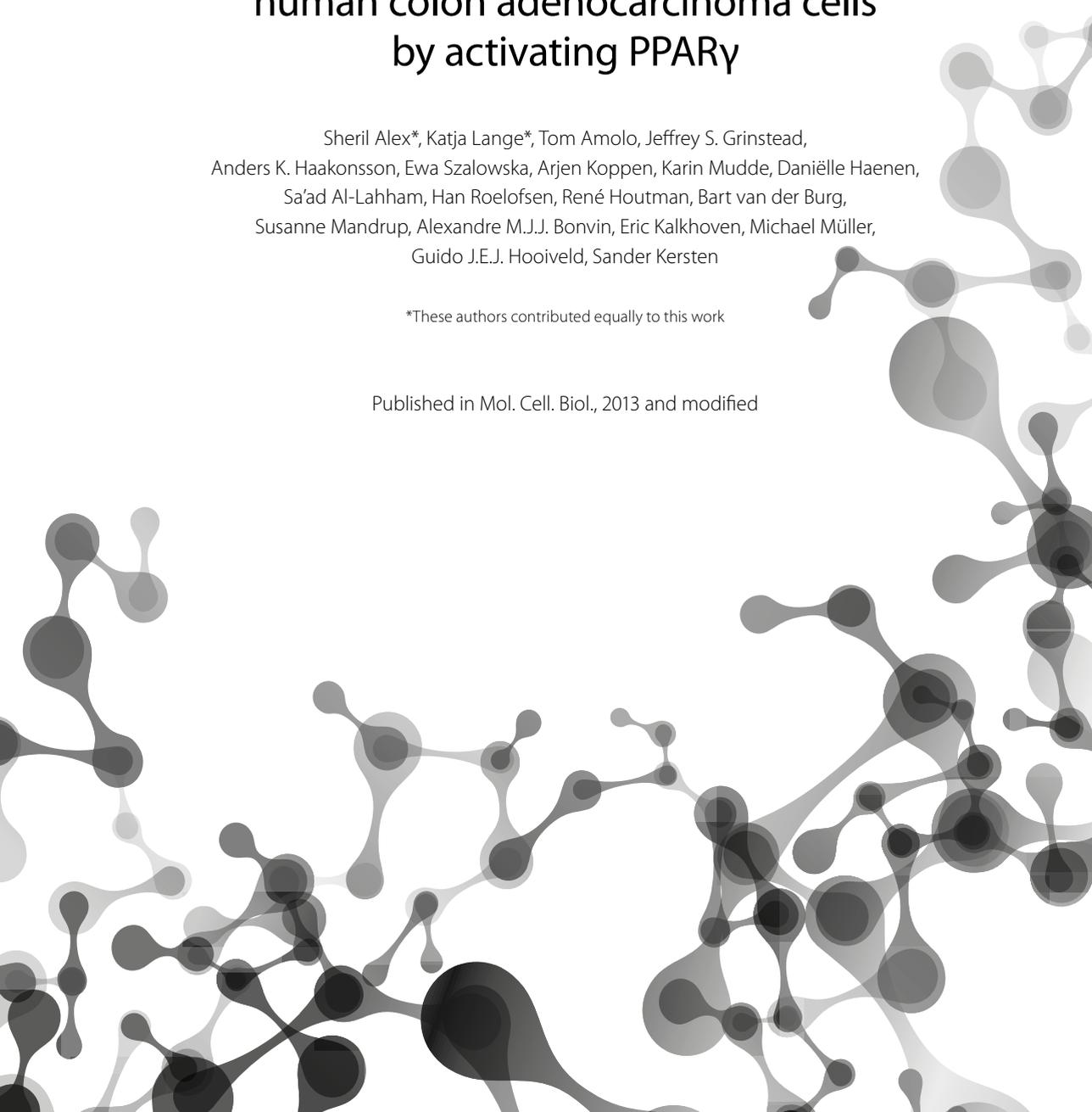
Chapter 5

Short chain fatty acids stimulate Angiopoietin-like 4 synthesis in human colon adenocarcinoma cells by activating PPAR γ

Sheril Alex*, Katja Lange*, Tom Amolo, Jeffrey S. Grinstead,
Anders K. Haakonsson, Ewa Szalowska, Arjen Koppen, Karin Mudde, Daniëlle Haenen,
Sa'ad Al-Lahham, Han Roelofsen, René Houtman, Bart van der Burg,
Susanne Mandrup, Alexandre M.J.J. Bonvin, Eric Kalkhoven, Michael Müller,
Guido J.E.J. Hooiveld, Sander Kersten

*These authors contributed equally to this work

Published in Mol. Cell. Biol., 2013 and modified



Abstract

Angiopoietin-like protein 4 (ANGPTL4/FAF) has been proposed as circulating mediator between the gut microbiota and fat storage. Here we show that transcription and secretion of ANGPTL4 in human T84 and HT-29 colon adenocarcinoma cells is highly induced by physiological concentrations of short chain fatty acids (SCFA). SCFA induce ANGPTL4 by activating the nuclear receptor PPAR γ , as demonstrated using PPAR γ antagonist, PPAR γ knock-down, and transactivation assays, which show activation of PPAR γ but not PPAR α and PPAR δ by SCFA. At concentrations required for PPAR γ activation and ANGPTL4 induction in colon adenocarcinoma cells, SCFA do not stimulate PPAR γ in mouse 3T3-L1 and human SGBS adipocytes, suggesting that SCFA act as selective PPAR γ modulators (SPPARM), which is supported by coactivator peptide recruitment assay and structural modelling. Consistent with the notion that fermentation leads to PPAR activation in vivo, feeding mice a diet rich in inulin induced PPAR target genes and pathways in the colon. We conclude that 1) SCFA potently stimulate ANGPTL4 synthesis in human colon adenocarcinoma cells; 2) SCFA transactivate and bind to PPAR γ . Our data point to activation of PPARs as a novel mechanism of gene regulation by SCFA in the colon, in addition to other mechanisms of action of SCFA.

Introduction

In the last decade, interest in the role of the intestinal microbiota has grown exponentially. The functional role of the gut microbiota has mainly been studied in relation to intestinal health and immune function. However, there are growing speculations that the gut microbiota may also influence other diseases, including type 1 diabetes, autism, and obesity, and potentially impact more distant organs [1]. Those notions have spurred the search for circulating factors that communicate between the intestinal microbiota and other parts of the body. One factor that was found to be strongly downregulated in intestine upon colonization of the gut of germ free mice with microbiota and that appears to be important for microbiota-induced deposition of triglycerides in adipocytes is Angiotensin-like 4 (ANGPTL4), also referred to as FIAF (Fasting Induced Adipose Factor)[2,3]. ANGPTL4 is secreted by a variety of tissues including adipose tissue, liver, skeletal muscle, and intestine [4]. It is released as a 50kD pro-hormone that is subsequently cleaved into N- and C-terminal fragments. The N-terminal fragment of ANGPTL4 blocks activity of the enzyme lipoprotein lipase (LPL), which catalyzes uptake of circulating lipids into tissues [5]. Accordingly, ANGPTL4 overexpression raises circulating triglyceride levels [6-10]. In cardiomyocytes and macrophages, induction of ANGPTL4 by fatty acids and subsequent inhibition of LPL is part of a feedback mechanism aimed at preventing cellular lipid overload and thus reducing lipotoxicity and inflammation [11,12].

Expression of *ANGPTL4* is under transcriptional control of Peroxisome Proliferator Activated Receptors (PPARs). Indeed, *ANGPTL4* was originally cloned as a target gene of PPAR α and PPAR γ [13,14]. Depending on the specific tissue, *ANGPTL4* mRNA levels are governed primarily by PPAR α (liver, small intestine)[13,15], PPAR δ (skeletal muscle, heart, macrophages) [11,12,16], or PPAR γ (adipocytes)[13,14]. In addition to the tissues mentioned above, *ANGPTL4* is also expressed in human colon [17](<http://biogps.org/>), but little is known about the factors regulating *ANGPTL4* mRNA expression and protein secretion in this tissue. PPAR δ is highly expressed in human colonocytes, followed by PPAR γ , and to a lesser extent PPAR α , suggesting that PPARs may be important regulators of *ANGPTL4* expression [18]. A similar expression profile is found in mouse colonocytes [19]. Considering the status of ANGPTL4 as putative mediator between the gut microbiota and adipose tissue, we were interested to identify the microbiota-related factors that influence ANGPTL4 production in the colon. Potential factors that are formed by intestinal microbial activity are short chain fatty acids, which are the most abundant ones in colon. Accordingly, we studied the regulation of ANGPTL4 by SCFA.



Materials and Methods

Animal experiments

Male C57BL/6J mice were purchased from Charles River Laboratories and housed in pairs on a 12-hour light/dark cycle at 21 °C with free access to feed and water throughout the entire experimental period. For 2 weeks before the start of the interventions, mice were fed a standard semi-synthetic low fat diet based on D12450B (Research Diets Inc., New Brunswick, NJ) with modified content of sucrose and corn starch (16.4% and 40.5% w/w, respectively). In the first experiment, 12-week-old mice were mildly sedated with isoflurane at 9:00 am. Mice were kept under sedation and received either a 80 µL rectal infusion of saline (n=4) or saline containing 100 mM sodium propionate (n=4). Both solutions had a pH of 6.5. The solutions were administered by inserting a gel loading tip 3 cm into the rectum and slowly pushing the solution out of the tip. The infusion were administered on 6 consecutive days. Four hours after the rectal infusion on day 6, mice were anaesthetized with isoflurane and subsequently sacrificed by cervical dislocation. The colon was excised and adhering fat was carefully removed. The epithelial lining of the colon was scraped and immediately frozen in liquid nitrogen to be stored at -80°C for subsequent RNA isolation. In the second experiment, mice were stratified according to their body weight and allocated to the control diet group (n=5, low fat diet as described above) or Inulin diet group (n=6, low fat diet with 10% w/w corn starch replaced by Inulin). After 10 days of dietary intervention, mice were fasted overnight and provided with 1 g of experimental diet. After 4 hours, mice were anesthetized using isoflurane and sacrificed by cervical dislocation. Colon was removed, luminal content and mucosal scrapings were collected, snap-frozen in liquid nitrogen and stored at -80°C (scrapings) or -20°C (content). Luminal content was used for SCFA measurement by gas chromatography. Mucosal scrapings were used for RNA isolation. The experiments were authorized by the Local Committee for Care and Use of Laboratory Animals at Wageningen University.

SCFA measurement

SCFA concentrations were measured in the luminal content of different parts of the gastrointestinal tract of C57BL/6 mice fed a low fat diet. After sacrifice, luminal content was collected in tubes containing phosphoric (H_3PO_4) and isocaproic acid, after which the tubes were thoroughly mixed and stored at -20°C. For analysis, samples were thawed, mixed on a vortex and centrifuged at 14.000 rpm for 5 min. The supernatant was collected and SCFA concentrations were determined by gas chromatography (Fisons HRGC Mega 2, CE Instruments, Milan, Italy) at 190°C using a glass column fitted with Chromosorb 101. The carrier gas was N_2 saturated with methanoic acid, and isocaproic acid was used as an internal standard.

Cell Culture

T84 and HT-29 human colonic carcinoma epithelial cells were cultured in DMEM F-12 or DMEM medium (Lonza, the Netherlands) containing 5% or 10% fetal calf serum (FCS), respectively, and 1% penicillin/streptomycin at 37°C in a humidified incubator of 95% air/5% CO₂. Cells were grown till 80-90% confluence and treated for the indicated time periods with sodium acetate, sodium propionate and sodium butyrate from a stock in PBS to a final concentration of 1 mM or 8 mM. Alternatively, cells were treated with synthetic agonists for PPAR α (Wy14643, 5 μ M), PPAR δ (GW501516, 1 μ M) and PPAR γ (rosiglitazone, 1 μ M) or synthetic PPAR γ antagonist GW9662 (5 μ M) or Trichostatin A (100 nM) from a stock in DMSO. In one experiment, cells were pre-incubated with α -amanitin (10 μ g/mL) from a stock in water for 1 hour followed by 24 hours co-incubation with butyrate at 1 mM. After the indicated time points medium was collected and used for analysis of ANGPTL4 by ELISA, and/or cells were harvested for subsequent RNA isolation.

siRNA transfection

Silencing of the PPAR γ gene in T84 cells was done using ON-TARGETplus SMARTpool for Human PPAR γ (Thermo Scientific Dharmacon, Etten-Leur, The Netherlands), representing a mixture of four siRNAs:CAAUACCAUUCGUUAUC,GACAUGAAUCCUUAUGA, GAUUAU-CAAGCCCUUCACUA, GACAGCGACUUGGCAAUAU. As control we used ON-TARGETplus Non-targeting Pool (Thermo Scientific Dharmacon). Transfection was done according to the manufacturers protocol using DharmaFECT 1 transfection reagent. The cells were transfected with 100 nM siRNA and incubated for 48 hours, followed by treatment with butyrate (1 mM) for a period of 24 hours. Thereafter, medium was collected for ANGPTL4 protein analysis by ELISA and the cells were used for RNA isolation and subsequent qPCR.

PPAR γ chromatin immuno-precipitation (ChIP)

T84 cells were grown in 10 cm plates to a density of approximately 5 million cells/dish and treated with butyrate (8 mM) or rosiglitazone (1 μ M). After 1 hour medium was removed and cells were fixed in 1% formaldehyde in PBS. After 10 minutes, crosslinking was stopped by the addition of glycine to a final concentration of 0.125 M. The cells were washed three times with ice cold PBS and re-suspended in 0.3 mL of lysis buffer (0.1% SDS, 1% TritonX-100, 0.15 M NaCl, 1mM EDTA, 20mM Tris pH 8.0). The lysate was frozen at -80°C until further analysis. ChIP was carried out as described in Siersbæk et al. [21], with the exceptions that disuccinimidyl glutarate was not used for cross-linking, and sonication was performed in 2x10 on/off cycles of 30 sec. The PPAR γ antibody used was PPAR γ H100 (sc-7196, Santa Cruz Biotechnologies, USA).

PPAR γ stable reporter assay

PPAR γ CALUX cell line was obtained from BioDetection Systems B.V. (Amsterdam, the Netherlands). PPAR γ CALUX cells are based on human osteoblastic (U2-OS) cells (American

Collection Cell Culture (ATCC), Manassas , VA, USA), stably transfected with a human PPAR γ expression plasmid and a luciferase reporter construct [22]. The cells were cultured as described previously in a 1:1 mixture of Dulbecco's modified Eagle's medium and Ham's F12 medium (Invitrogen, Breda, the Netherlands) supplemented with 7.5% fetal bovine serum (FBS) (Invitrogen), 1% nonessential amino acids (Invitrogen) and penicillin/streptomycin (Invitrogen) to final concentrations 10U/mL and 10 μ g/mL, respectively [22,23]. Once per month 200 μ g/mL G418-disulfate was added to the culture medium.

The PPAR γ CALUX assay was performed as described previously [22,24]. In short, for the exposure experiments PPAR γ -CALUX cells were plated in the 96-well plates with phenol-free DF supplemented with 5% dextran coated charcoal-stripped FBS (Invitrogen) at a volume of 200 μ L per well. The next day medium was refreshed and cells were incubated in triplicates with test compounds added to the culture medium. After 24 hours, the cells were checked visually for cytotoxicity, the medium was removed, and the cells were lysed in Luciferase Cell Culture Lysis Reagent (Promega, Leiden, the Netherlands). Luciferase activity was measured in cellular extracts using a Synergy HT Multi-Detection Microplate Reader (Luminometer of BioTek Instruments Inc., USA). For each test compound at last two independent experiments were performed.

GAL4-PPAR stable reporter assay

The assay was carried out in Hela cells stably expressing a chimeric protein containing the ligand binding domain (LBD) of human PPAR α , PPAR δ or PPAR γ fused to the yeast transactivator GAL4 DNA binding domain (DBD) and a luciferase (Luc) reporter gene driven by a pentamer of the GAL4 recognition sequence in front of a β -globin promoter. The cells were seeded in 96-well plates and incubated the following day with tested compounds for 24h. At the end of the incubation, the luciferase activity was measured using a BMG LUMIstar Galaxy luminometer. Results expressed as relative light units (RLU) were obtained from experiments performed in triplicate. Rosiglitazone , GW7647 and L165041 (all at 1 μ M) were used as positive controls for activation of PPAR α , PPAR δ and PPAR γ , respectively. The measurements were performed as a commercial service by Tebu-bio laboratories (http://www.tebu-bio.com/upload/cms/-PPAR_Screening_Poster.pdf), Le Perray-en-Yvelines, France.

ANGPTL4 ELISA

ANGPTL4 levels in the medium were measured by ELISA as detailed previously [17]. Briefly, 96-well plates were coated with anti-human ANGPTL4 polyclonal goat IgG antibody (AF3485, R&D Systems) and incubated overnight at 4°C. Plates were washed extensively between each step. After blocking, 100 μ L of medium of cells was applied, followed by 2 hour incubation at room temperature. A standard curve was prepared using recombinant human ANGPTL4 (3485-AN, R&D Systems) at 0.3 to 2.1 ng/well. Next, 100 μ L of diluted biotinylated anti-human ANGPTL4 polyclonal goat IgG antibody (BAF3485, R&D Systems)

was added for 2 hour, followed by addition of streptavidin-conjugated horseradish peroxidase for 20 min, and tetramethyl benzidine substrate reagent for 6 min. The reaction was stopped by addition of 50 μ L of 10% H₂SO₄, and the absorbance was measured at 450 nm.

Cofactor recruitment assay

Nuclear receptor PamChip arrays (PamGene, s'Hertogenbosch, The Netherlands) were used as described previously [25]. Upon binding a ligand, PPARs undergo a conformational change that promotes the formation of a cofactor binding pocket, subsequently allowing interaction with the so-called LxxLL motif within some coregulators. The PamChip arrays consist of 53 peptides encompassing the LxxLL motifs of 21 different coregulator proteins. The sequences are provided as supplementary data in Koppen et al. [25]. Briefly, the arrays were incubated with glutathione S-transferase-tagged PPAR γ ligand binding domain (LBD) (Invitrogen, Breda, The Netherlands) in the presence and absence of the ligands i.e. sodium acetate, sodium propionate, sodium butyrate (each at 40 mM), and rosiglitazone (1 μ M). Quantification of the interaction between PPAR γ and coregulators was made using Alexa 488-conjugated anti-glutathione S-transferase rabbit polyclonal antibody (Invitrogen).

Affymetrix GeneChip microarray analysis

Microarray analysis was performed on T84 cells treated with 0.1 μ M rosiglitazone or 1 mM butyrate for 24h. Total RNA was prepared using TRIzol reagent (Invitrogen), and purified on columns using the RNeasy Mini Kit (Qiagen, Venlo, The Netherlands). RNA quality and integrity were verified with the RNA 6000 Nano assay on the Agilent 2100 Bioanalyzer (Agilent Technologies, Amsterdam, the Netherlands). Hybridization, washing, and scanning of the Affymetrix Human Gene 1.1 ST Array Plate was performed on Affymetrix GeneTitan. Scans of the Affymetrix arrays were processed using packages from the Bioconductor project [26]. Raw signal intensities obtained by robust multiarray (RMA) normalization [27]. Probe-sets were defined according to Dai et al. using remapped chip definition file (CDF) version 15 based on the Entrez gene database [28]. The microarray data were submitted to Gene Expression Omnibus (Accession GSE40706).

Microarray analysis of colon of mice fed inulin was carried out as described above using the Affymetrix Mouse Gene 1.1 ST Array Plate (Accession GSE43065).

3T3-L1 adipogenesis

3T3-L1 fibroblasts were amplified in DMEM/10% calf serum and subsequently seeded into six-well plates. Two days after the cells reached confluence, the medium was changed to DMEM/10% fetal calf serum containing 0.5 mM 3-isobutyl-1-methylxanthine, 2 μ g/ml insulin (Actrapid), 0.5 μ M dexamethasone, and either the short chain fatty acids or rosiglitazone (1 μ M). After two days, the medium was changed to DMEM/10% fetal calf

serum containing 2 µg/ml insulin and the short chain fatty acids or rosiglitazone. Cells were harvested after two more days for RNA isolation (day 4). This mild adipogenesis protocol permits assessment of PPAR γ agonist activity of added compounds [29]. Oil Red O staining was performed at day 10 using a standard protocol.

In a second protocol, two days after the cells confluence (=day 0), the medium was changed to DMEM/10% fetal calf serum containing 0.5 mM 3-isobutyl-1-methylxanthine, 5 µg/ml insulin, and 1 µM dexamethasone. After three days, the medium was changed to DMEM/10% fetal calf serum containing 5 µg/ml insulin. The medium was subsequently changed every 3 days, and no further insulin was added after 6 days. At day 10, 3T3-L1 adipocytes were treated with rosiglitazone (1 µM) and the short chain fatty acids (8 mM) for 24 hours.

Adipogenesis in SGBS

Human Simpson–Golabi–Behmel syndrome (SGBS) adipocytes were cultured and grown to confluency in 0F medium (DMEM/F12/Biotin-panthotenate/PSA) plus 10% FCS. One day post-confluency, the medium was changed to Quick-diff medium which is 3FC medium (0F medium plus 0.01 mg/mL human apo-transferrin, 2×10^{-8} M insulin, 10^{-7} M cortisol, 0.2 nM T3) with 50 nM dexamethasone, 500 µM 3-isobutyl-1-methylxanthine and 2 µM Rosi or 8 mM SCFA. On day 4, medium was changed to 3FC medium and changed every 4 days until day 15. Expression of differentiation markers and PPAR γ targets was determined by qPCR at day 15. DMSO was used as control.

RNA isolation and qPCR

RNA was isolated from T84 cells using RNeasy columns or Trizol. 1 µg of total RNA was reverse-transcribed with iScript (Bio-Rad, Veenendaal, the Netherlands) or Fermentas cDNA synthesis kit according to the instructions from the manufacturer. cDNA was amplified on a Bio-Rad CFX384 Real Time System using Sensimix (Bioline, GC Biotech, Alphen aan de Rijn, The Netherlands). Cyclophilin or 36B4 was used as housekeeping gene. PCR primer sequences were taken from the PrimerBank and ordered from Eurogentec (Seraing, Belgium). Primer sequences are presented in table 1.

Modeling of Butyrate Binding to PPAR γ

For modeling the structure of the complex between PPAR γ and butyrate, HADDOCK version 2.1 was used [30]. HADDOCK is a highly successful modeling approach that makes use of structural knowledge when available to drive the docking procedure. In this case, the x-ray crystal structure of PPAR γ with decanoic acid bound (3U9Q) was used to identify the likely binding site for short chain fatty acids such as butyrate [31]. The docking was performed using the web server version of HADDOCK [32], with the following modifications. The high temperature rigid body torsion angle dynamics and the first rigid body cooling stage were not performed (number of steps set to 0) and the second cooling phase of

Table 1 list of primers used

Gene	Forward primer	Reverse primer
m36B4	ATGGGTACAAGCGTCCTG	GCCTTGACCTTTTCAGTAAG
Fabp4	AAGAAGTGGGAGTGGGCTTT	AATCCCCATTACGCTGATG
Gyk	ATCCGCTGGCTAAGAGACAACC	TGCACTGGGCTCCAATAAGG
Slc2a4	GGAAGGAAAAGGGCTATGCTG	TGAGGAACCGTCCAAGAATGA
Adipoq	GCAGAGATGGCACTCCTGGA	CCCTTCAGCTCCTGTCATTCC
Angptl4	GTTTGCAGACTCAGCTCAAGG	CCAAGAGGTCTATCTGGCTCTG
ANGPTL4	CACAGCCTGCAGACACAACCTC	GGAGGCCAAACTGGCTTTGC
PPARG	GAGCCCAAGTTTGAGTTTGC	CAGGGCTTGTAGCAGTTGT
PPARA	CAGAACAAGGAGGCGGAGGTC	TTCAGGTCCAAGTTTGCGAAGC
PPARD	TGGCTTTGTCACCCGTGAGT	ACAGAATGATGGCCGCAATGAA
PLIN2	ATGGCATCCGTTGCAGTTGAT	GATGGTCTTCACACCGTTCTC
UCP2	TGCCCTCTGAAAGCCAAC	CTTGACCACGTCTACAGGGGA
AQP8	GCGAGTGCCTGGTACGAAC	CAGGCACCCGATGAAGATGAA
LGALS1	TCGCCAGCAACTGAATCTC	GCACGAAGCTCTTAGCGTCA
h36B4	CGGGAAGGCTGTGGTGCTG	GTGAACACAAAGCCACATTCC

torsion angle dynamics with flexible side-chains at the interface was started with a reduced initial temperature of 500K instead of the default 1000K.

All calculations were performed with CNS1.2 [33]. Non-bonded interactions were calculated with the OPLS force field using a cutoff of 8.5 Å [34]. The electrostatic potential (Eelec) was calculated by using a shift function, while a switching function (between 6.5 and 8.5Å) was used to define the Van der Waals potential (EvdW). The HADDOCK score was used to rank the generated poses. It is a weighted sum of intermolecular electrostatic (Eelec), van der Waals (EvdW), desolvation (Edesolv) and ambiguous interaction restraint (AIR) energies with weight factors of 0.2, 1.0, 1.0 and 0.1, respectively.

The ambiguous interaction restraints used to drive the docking of butyrate to PPAR γ were derived from the X-ray crystal structure of PPAR γ with decanoic acid (DA) bound (3U9Q), where protein residues near the decanoic acid were identified as the butyrate binding site. All protein residues within 3.9Å of the decanoic acid (residues 278, 281, 282, 285, 289, 323, 356, 360, 363, 449, 469, and 473) were defined as active restraints for the rigid body docking phase, and as passive restraints for the subsequent semi-flexible refinement stage. The ligand was considered active for both docking phases. This strategy effectively pulls the butyrate ligand into the binding site during rigid-body docking while allowing a thorough exploration of the binding pocket during the refinement stage. With the protein

active site residues defined as passive during the semi-flexible refinement stage, the ligand is not restrained to any particular orientation or location within the defined binding site. For comparison, a second docking trial was performed with active residues for the entire PPAR γ binding site, defined as amino acids within 3.9Å of either the partial agonist decanoic acid (from 3U9Q, residues 278, 281, 282, 285, 289, 323, 356, 360, 363, 449, 469, and 473) or the full agonist nitrosylated fatty acid (from 3CWD, residues 285, 286, 288, 289, 326, 327, 330, 340, 341, 364, 449, and 473). Docking with this wider binding pocket definition did not give additional solutions.

At the end of the docking protocol, clustering based on ligand pairwise root mean square deviation criteria was performed and the best scoring structure of the best scoring cluster was taken as the best solution. The clustering distance cutoff was reduced from the default 7.5Å cutoff for protein-protein complexes to 1.75Å, which is more suitable for protein-ligand complexes.

The interactions between decanoic acid and PPAR γ were analyzed by submitting the X-ray crystal structure 3U9Q to the HADDOCK refinement server [32] which only performs a gentle refinement in explicit solvent and returns similar statistics as a full docking run.

Results

ANGPTL4 production is stimulated by short chain fatty acids

One set of molecules that may mediate the effect of microbiota on *ANGPTL4* expression in the colon are the short chain fatty acids (SCFA) acetate, propionate and butyrate. To estimate the concentration of fatty acids to be used for *in vitro* experiments, we determined the total SCFA concentration in the lumen of different parts of the gastrointestinal (GI) tract of mice fed a low fat diet (Figure 1A). Total SCFA concentration varied from <10 mM in the proximal small intestine to >40 mM in the caecum. In the colon the highest concentrations were found for acetate, followed by propionate and butyrate (Figure 1B). In the literature, even higher concentrations have been reported that exceed 100 mM [35]. Based on these data, we used low to medium millimolar SCFA concentrations.

To study regulation of *ANGPTL4* production in colonocytes, we used the human colonic adenocarcinoma cell lines T84 and HT29. Remarkably, all SCFA significantly increased *ANGPTL4* secretion by T84 and HT-29 cells (Figure 1C). Strongest effects were observed for butyrate, followed by propionate and acetate. Increased *ANGPTL4* secretion was mirrored by a pronounced increase in *ANGPTL4* mRNA (Figure 1D), suggesting SCFA stimulate *ANGPTL4* transcription, which was further supported by the rapid time-course of induction of *ANGPTL4* mRNA by butyrate (Figure 1E). Furthermore, induction of *ANGPTL4* secretion by butyrate was almost completely abolished by α -amanitin, an inhibitor of RNA polymerase II (Figure 1F). In support of *in vivo* regulation of *Angptl4* gene expression by SCFA, rectal infusion of propionate significantly increased *Angptl4* mRNA expression in mouse colon (Figure 1G).

Butyrate and to a lesser extent propionate are able to inhibit histone deacetylase (HDAC) activity, while acetate is ineffective [36,37]. To study the impact of HDAC inhibition on ANGPTL4 secretion, we treated T84 and HT-29 cells with the specific HDAC inhibitor trichostatin A, which similar to butyrate inhibits class I and class II HDACs. Unlike butyrate, trichostatin A had no effect on ANGPTL4 secretion (Figure 1H), indicating that the stimulatory effect of butyrate on ANGPTL4 is likely independent of HDAC inhibition. The effectiveness of trichostatin A was demonstrated by the marked induction of the LGALS1 gene (Figure 1I).

PPAR γ mediates induction of ANGPTL4 by SCFA

Because the medium chain fatty acid decanoic acid was recently shown to be a direct ligand of PPAR γ and ANGPTL4 is a sensitive target of PPAR γ in colonocytes [31,38], we hypothesized that butyrate may upregulate *ANGPTL4* expression in colonocytes by activating PPAR γ . To test the responsiveness of ANGPTL4 to PPAR activation in T84 and HT-29 cells, the cells were treated with synthetic agonists for PPAR α , PPAR δ and PPAR γ . The PPAR γ agonist rosiglitazone was the most potent inducer of ANGPTL4 protein secretion (Figure 2A) and mRNA (Figure 2B) in both cell types, followed by PPAR δ agonist GW501516 and PPAR α agonist Wy14643. These results are consistent with the high expression of *PPARG* mRNA in T84 cells (Figure 2C, top panel) and to a lesser extent in HT-29 cells (lower panel), and the corresponding protein levels of the three PPARs in the two cell lines [39-43]. Upon incubation with rosiglitazone, a time- and dose-dependent increase in ANGPTL4 protein was observed in T84 cells (Figure 2D,E), indicating that ANGPTL4 production is highly sensitive to PPAR γ activation. Accordingly, we considered PPAR γ as a possible candidate mediating induction of *ANGPTL4* mRNA by butyrate.

To investigate whether the stimulatory effect of butyrate on ANGPTL4 secretion occurs via PPAR γ , we used the PPAR γ antagonist GW9662. Induction of ANGPTL4 secretion by butyrate in T84 cells was strongly diminished by GW9662 (Figure 3A), suggesting that butyrate increases ANGPTL4 secretion mainly by activating PPAR γ . Similarly, siRNA mediated knock-down of PPAR γ mRNA by about 60% led to a very significant 40% decrease in butyrate-induced *ANGPTL4* mRNA level (Figure 3B) and butyrate-induced ANGPTL4 protein secretion (Figure 3C), again suggesting that the effect of butyrate on ANGPTL4 is largely mediated by PPAR γ . Stimulation of PPAR γ mRNA by butyrate (Figure 3B) likely reflects PPAR γ autoregulation. To further pursue the role of PPAR γ in ANGPTL4 regulation by butyrate, we performed chromatin immunoprecipitation to study binding of PPAR γ to the *ANGPTL4* gene. The results indicate that PPAR γ occupies the PPAR binding site within intron 3 of the *ANGPTL4* gene (Figure 3D), which is known to mediate PPAR responsiveness [44,45]. Occupancy was modestly induced by both butyrate and rosiglitazone. Taken together, these data strongly suggest that induction of ANGPTL4 by butyrate in T84 cells is mediated by PPAR γ .

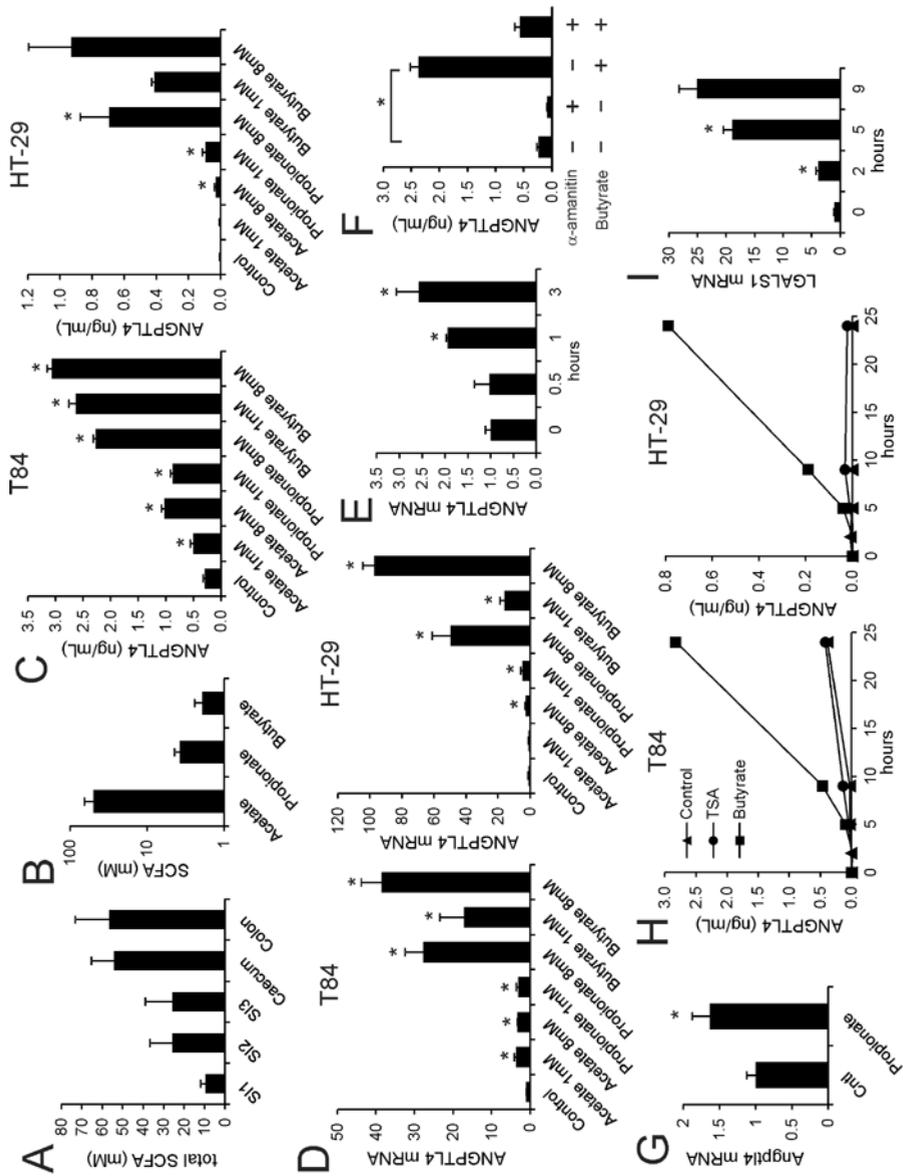


Figure 1 Physiological concentrations of SCFA stimulate ANGPTL4 mRNA and protein secretion in colon adenocarcinoma cells

A) Total SCFA concentration in different sections of the intestine of mice fed low fat diet. Errors bars represent SEM. B) Concentration of individual SCFA in colon of mice fed low fat diet. Errors bars

represent SEM. C) ANGPTL4 concentration in medium of T84 and HT-29 cells treated with SCFA for 24 hours at the indicated concentrations. D) ANGPTL4 mRNA in T84 and HT-29 cells treated with SCFA for 24 hours at the indicated concentrations. E) Time-course of induction of ANGPTL4 mRNA in T84 cells by butyrate (1 mM). F) Inhibitory effect of α -Amanitin (10 μ g/mL) on induction of ANGPTL4 secretion by 1 mM butyrate in T84 cells. G) Effect of rectal infusion of propionate on Angptl4 mRNA in epithelial scrapings of mouse colon. Errors bars represent SEM. H) Time-course of regulation of ANGPTL4 protein in medium of T84 and HT-29 cells by trichostatin A (100 nM) and butyrate (8 mM). I) Stimulatory effect of trichostatin A (100 nM) on LGALS1 mRNA in T84 cells. Errors bars represent SD except when indicated otherwise. Asterisk indicates significantly different from control according to Student's t-test ($p < 0.05$).

To investigate whether butyrate acts as general inducer of PPAR γ -dependent gene regulation, we performed microarray on T84 cells treated either with butyrate (1 mM and 8 mM) or rosiglitazone. Remarkably, effects of butyrate on global gene expression were much more pronounced compared to rosiglitazone, likely linked to its property as potent HDAC inhibitor (Figure 4A, Suppl. Figure 1). Interestingly, more overlap was observed between 1 mM butyrate and rosiglitazone than between 8 mM butyrate and rosiglitazone (Suppl. Figure 1). Cluster analysis identified clusters of genes induced by rosiglitazone and butyrate (Figure 4A). Genes in one cluster, which besides *ANGPTL4* included PPAR targets *HMOX1*, *PDK4* and *UCP2* [46-48], were induced by 1 mM and 8 mM butyrate. Genes in another cluster, which included other PPAR targets including *AQP8* and *PLIN2* [49,50], were induced by 1 mM butyrate and surprisingly not induced or even suppressed by 8 mM butyrate. The role of PPAR γ in mediating the stimulatory effect of 1 mM butyrate on expression of *PLIN2* and *UCP2*, representing two distinct clusters, was verified by siRNA (Figure 4B).

The biphasic regulation of *AQP8* and *PLIN2* by butyrate but not by propionate and acetate was confirmed by qPCR (Figure 4C). Such a biphasic regulation is only possible if a gene is regulated by at least two distinct mechanisms: a stimulatory effect presumably via PPAR γ already active at lower butyrate concentrations and a suppressive effect impacting a subset of PPAR targets at higher butyrate concentrations. Overall, these data indicate that SCFA act as a general inducer of PPAR γ -dependent gene regulation in T84 cells. However, especially at higher concentrations, the main effects of SCFA in T84 cells are independent of PPAR γ .

SCFA function as PPAR γ agonists

To further study activation of PPAR γ by SCFA, we employed a stable PPAR γ reporter assay. Hela cells stably transfected with a fusion construct between the DNA-binding domain of Gal4 and the ligand-binding domain of PPAR γ , PPAR α , or PPAR δ were incubated with increasing concentrations of SCFA. Remarkably, PPAR γ was strongly activated by butyrate and to a lesser extent by propionate (Figure 5A). In contrast, PPAR α and PPAR δ were not activated by any of the SCFA, even at higher concentrations.

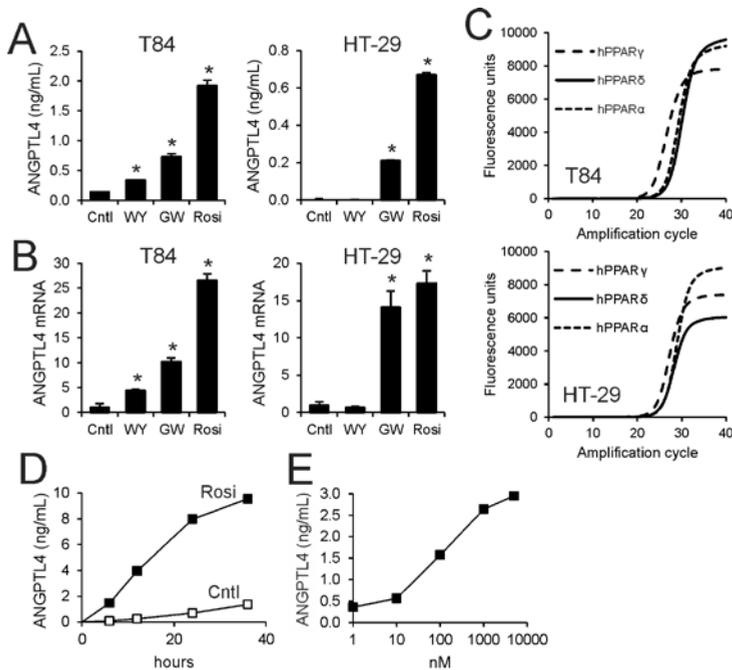


Figure 2 PPAR γ potently stimulates ANGPTL4 in colon adenocarcinoma cells

Synthetic agonists for PPAR α (Wy14643, 5 μ M), PPAR δ (GW501516, 1 μ M) and PPAR γ (rosiglitazone, 1 μ M) stimulate ANGPTL4 secretion (A) and mRNA expression (B) in T84 and HT-29 cells. C) Amplification curve of PPAR α , PPAR δ , and PPAR γ mRNA as determined by qPCR in T84 and HT-29 cells. Size of amplicons varied less than 10%. D) Time course of induction of ANGPTL4 protein in medium by rosiglitazone (1 μ M). E) Dose-dependent induction of ANGPTL4 protein in medium by rosiglitazone. Unless indicated otherwise, cells were treated for 24h. Errors bars represent SD. Asterisk indicates significantly different from control according to Student's t-test ($p < 0.05$).

To corroborate activation of PPAR γ by SCFA we performed an alternative reporter assay in U2OS cells stably transfected with a human PPAR γ expression and a PPRE-luciferase reporter construct. Similar to the results described above, propionate but especially butyrate markedly activated PPAR γ reporter activity, while acetate had minimal effect (Figure 5B). The threshold for stimulating reporter activity was 0.5 mM for butyrate compared to 10 nM for rosiglitazone, indicating that butyrate has relatively weak PPAR γ agonist activity.

To explore the possibility that SCFA serve as direct PPAR γ agonists, we used Nuclear Receptor PamChip[®] Arrays. In this system the interaction between soluble nuclear receptors and 53 immobilized peptides corresponding to specific coregulator-nuclear receptor binding

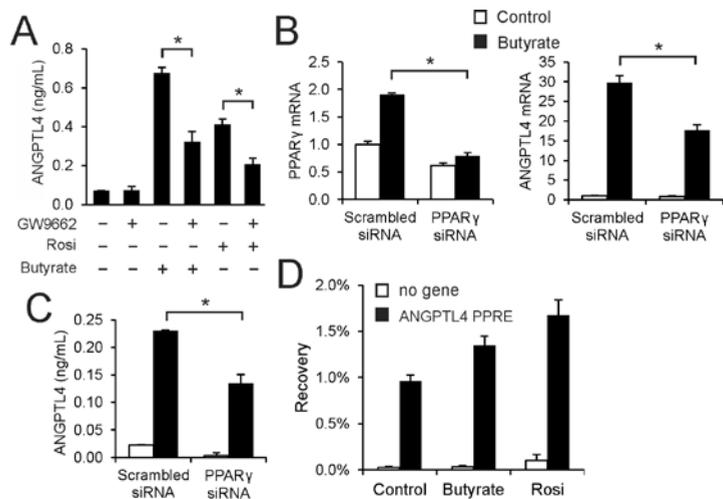


Figure 3 Induction of ANGPTL4 by butyrate is mediated by PPAR γ

A) Inhibitory effect of PPAR γ antagonist GW9662 (5 μ M) on induction of ANGPTL4 secretion in medium by rosiglitazone (10 nM) and butyrate (1 mM) in T84 cells. Effect of siRNA mediated PPAR γ knock-down (B, left panel) in T84 cells on butyrate-induced upregulation of ANGPTL4 mRNA (B, right panel) and ANGPTL4 protein in medium (C). D) PPAR γ ChIP-qPCR on selected loci in T84 cells treated with butyrate (8 mM) or rosiglitazone (1 μ M) for 24 hours. Bars represent the mean recovery plus range of two independent experiments. Errors bars represent SD except when indicated otherwise. Asterisk indicates significantly different from control according to Student's t-test ($p < 0.05$).

regions is studied. Using this system for PPAR γ , we previously generated ligand-specific coregulator interaction profiles [25]. Both rosiglitazone and butyrate promoted the interaction between PPAR γ and numerous coactivator peptides (e.g. CBP; Figure 5C). However, in contrast to rosiglitazone, butyrate did not relieve the interaction between PPAR γ and NCoR1 and NCoR2 corepressor peptides. Consistent with the PPAR γ reporter data, acetate had little effect. Quantitative analysis of the PamChip[®] Arrays is presented in Figure 5D. These data suggest that butyrate functions as selective PPAR γ modulator (SPPARM) [51].

Modeling of butyrate bound to PPAR γ

To find support for SCFA behaving as selective PPAR γ modulators we performed structural modeling of the binding of butyrate to PPAR γ using HADDOCK [30,32] (Figure 6A). Docking of butyrate into PPAR γ resulted in three different clusters of solutions, each showing butyrate bound in the ligand-binding pocket of the protein. The best cluster in terms of HADDOCK score was also the largest one (179/200 calculated structures) (Table 2) and

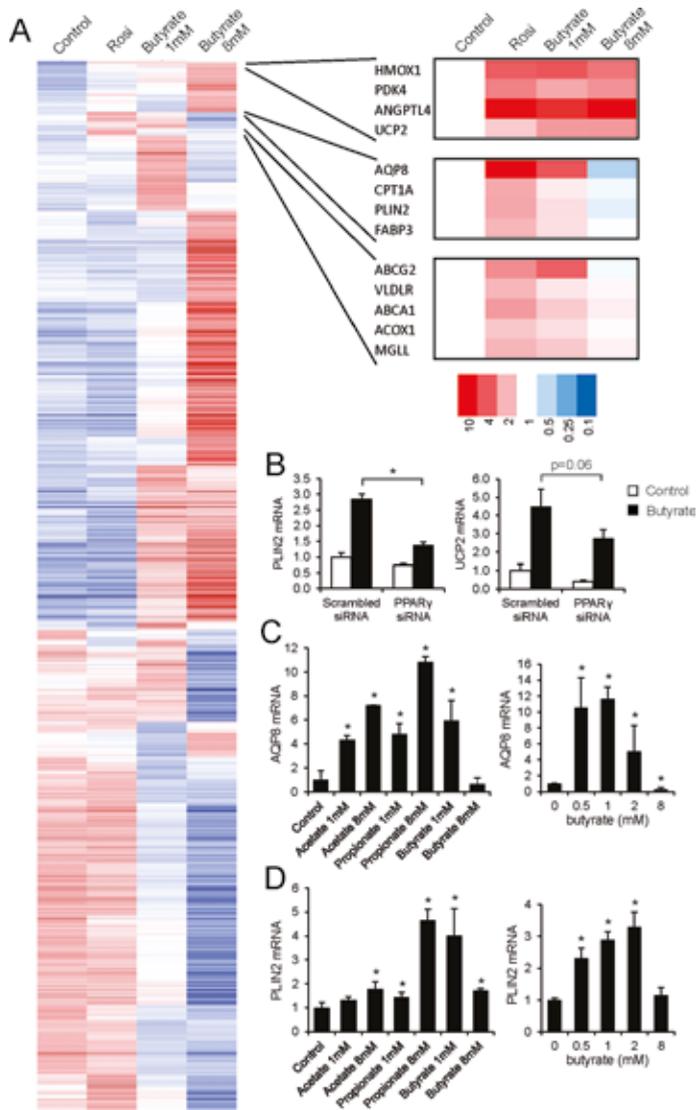


Figure 4 Global effects of butyrate in T84 cells

T84 cells were treated with rosiglitazone (1 μ M) or butyrate (1 mM and 8 mM) for 24 hours and subjected to Affymetrix microarray analysis. A) hierarchical clustering based on Pearson's correlation with average linkage. B) Effect of siRNA mediated PPAR γ knock-down in T84 cells on induction of PLIN2 and UCP2 mRNA by butyrate. AQP8 mRNA (C) and PLIN2 mRNA (D) levels were determined in T84 cells treated with SCFA for 24 hours at the indicated concentrations. Errors bars represent SD. Asterisk indicates significantly different from control according to Student's t-test ($p < 0.05$).

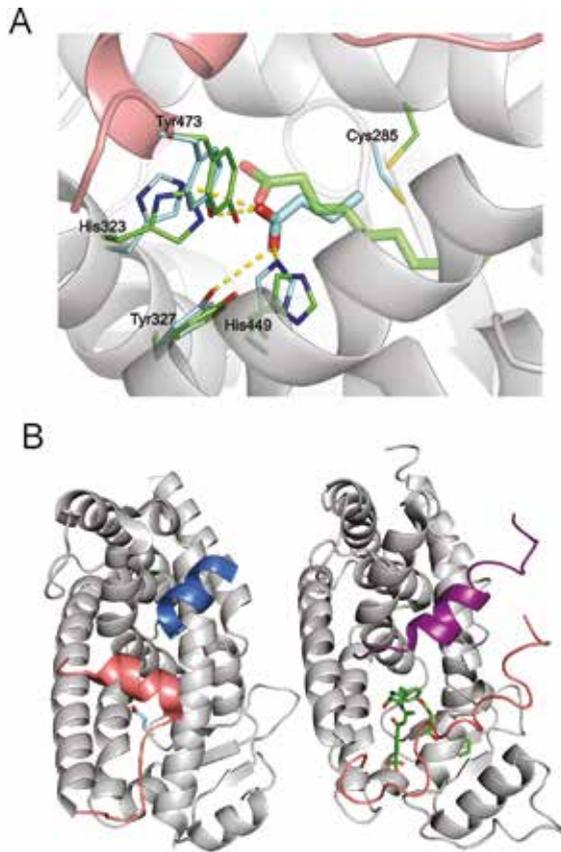


Figure 6 Modeling of butyrate into the PPAR γ binding pocket

A) The model reveals the complex between butyrate and PPAR γ with the best HADDOCK score (butyrate shown as cyan sticks), overlaid with the crystal structure of the decanoic acid complex with PPAR γ (3U9Q, decanoic acid shown as green sticks) by aligning the protein backbone atoms of the two structures (ribbon displayed for the HADDOCK model). The displayed protein side-chains are shown as thin cyan or green sticks, and the sidechains making contacts with the docked butyrate or decanoic acid, respectively. Hydrogen bond contacts between the butyrate and the protein are shown as yellow dashed lines. B) Comparison of the binding location for coactivator peptide PGC-1 α (blue ribbon) bound to PPAR γ (HADDOCK model of the PPAR γ -butyrate complex, left) with the binding site for the SMRT corepressor peptide (purple ribbon) binding to PPAR α (1KKQ, right) [69]. In both structures the C-terminal portion of the PPAR molecule that forms helix AF-2 is colored pink for comparison. The structures were aligned using the backbone atoms of the receptors.

Table 2 Comparison of PPAR γ complexes with fatty acids^a: 3U9Q is the HADDOCK-refined crystal structures of the PPAR γ -decanoic acid complex, while the other three clusters were obtained by HADDOCK modeling of the butyrate – PPAR γ complex

Complex	BSA (Å ²)	HADDOCK score ^b [a.u.]	Ligand-RMSD ^c (Å)	Size of cluster
3U9Q	478 ± 6	-22.1 ± 2.9	0.0	n/a
Cluster 1-butyrate	286 ± 14	-7 ± 3	1.55 ± 0.29	179
Cluster 2-butyrate	280 ± 6	0.1 ± 4	4.70 ± 0.77	13
Cluster 3-butyrate	275 ± 6	4.6 ± 4	0.99 0.13	6

a) Averages and standard deviations calculated over the best four members of a set of structures are reported.

b) HADDOCK score as defined in Materials and Methods.

c) Positional RMSD of the ligand atoms calculated after superimposition of the protein backbone atoms.

contained the best scoring models. It reveals a binding mode very similar to decanoic acid within the crystal structure of PPAR γ (3U9Q) [31]. Specifically, the average RMSD between the butyrate and decanoic acid common atoms upon aligning the backbone atoms of the PPAR γ (ligand-RMSD) was 1.55±0.29Å for the best four cluster members. The remaining clusters have less favorable interaction energy.

Interactions between butyrate and PPAR γ are dominated by electrostatic interactions and hydrogen bonds between the carboxylate group and several hydrogen bond donating protein side-chains, which are very similar to those found in the structure with decanoic acid (Figure 6A). In the latter, hydrogen bonds are found to four different side-chains near the carboxylate group, namely His323 HE2, Ser289 OG, His449 HE2, and Tyr473 OH. The top four structures of the best cluster reveal specific hydrogen bonds from the butyrate oxygens to His323, Tyr473, and His449. The hydrogen bond to Ser289 is never observed, while in two of the structures a fourth hydrogen bond is made to Tyr327 OH. The average buried surface area for the best four docking solutions is 286 ± 14 Å², and the average HADDOCK score for the best four structures is -7 ± 3. In comparison, the buried surface area and HADDOCK score calculated for the decanoic acid-PPAR γ structure are 478Å², and -22.1, respectively.

We next explored the possible influence of butyrate on coactivator and corepressor binding (Figure 6B). Fortunately, the structure of PPAR γ with decanoic acid also contained a bound PGC1 α coactivator peptide, allowing us to model butyrate docked within the crystal structure of PPAR γ complexed with the PGC-1 α peptide. However, no structures were available for co-repressor (peptides) binding to PPAR γ , which limited us to the crystal structure of PPAR α complexed with an antagonist and the SMRT corepressor peptide (1KKQ). As shown in Figure 6B, the binding sites for the coactivator (left, shown in blue) and corepressor (right, shown in purple) peptides occupy structurally analogous positions

on PPAR γ and PPAR α , and are likely to be mutually exclusive binding events. The C-terminal AF2 helix of PPAR γ (shown in pink) forms part of the binding site for the coactivator peptide PGC-1 α , and stabilization of this helix is known to promote binding of coactivator peptides [31]. As can be seen from the structural comparison, the AF2 helix of PPAR γ also occludes part of the structurally analogous binding site for the corepressor peptide binding to PPAR α . The butyrate molecule in the HADDOCK model makes several contacts to the AF2 helix (such as the hydrogen bond to Tyr473 OH) that are very similar to the contacts that were observed between decanoic acid and the AF2 helix in the X-ray crystal structure of the decanoic acid-PPAR γ complex (3U9Q), which is consistent with our finding using coactivator peptide binding assay that butyrate can promote at least partial agonism of the PPAR γ receptor, but does not promote dissociation of corepressor peptides.

SCFA do not induce PPAR γ -mediated adipogenesis

Given that SCFA can activate PPAR γ -mediated transcription in reporter assays but exhibit only a partial overlap with a full PPAR γ agonist in their coregulator binding profile, SCFA may not display the same capacity as full agonists in cell-based assays. Since PPAR γ is a critical regulator of adipogenesis and mediates stimulation of adipogenesis by rosiglitazone, we subsequently used stimulation of 3T3-L1 adipogenesis as in vitro readout for PPAR γ activation by SCFA. It has been shown that NCoR1 represses 3T3-L1 adipogenesis and that dismissal of NCoR1 is essential for induction of adipogenesis by PPAR γ agonists [52,53]. Since SCFA/butyrate did not release the association between NCoR1 and PPAR γ as observed using the PamChip array, it may be expected that SCFA are unable to mimic the adipogenic effects of PPAR γ agonists [29]. Consistent with this notion and in contrast to rosiglitazone, SCFA did not induce 3T3-L1 adipogenesis, as revealed by oil red O staining (Figure 7A) and gene expression analysis of adipogenesis markers, which also represent PPAR γ targets (Figure 7B). In fact, at higher concentrations butyrate and propionate suppressed adipogenesis (Figure 7C). Furthermore, in contrast to rosiglitazone, butyrate did not upregulate PPAR γ targets in differentiated 3T3-L1 adipocytes (Figure 7D). Similarly, in contrast to rosiglitazone, SCFA were unable to induce adipogenesis in the human Simpson–Golabi–Behmel syndrome (SGBS) adipocyte model, as revealed by lack of induction of adipogenesis markers (Figure 7E). Together with the data from colon cell lines and PamChip® Arrays, these data support the classification of SCFA as selective PPAR γ modulators.

Consumption of inulin leads to PPAR activation in mouse colon

To find support for activation of PPAR γ by SCFA in vivo, we studied the effect of dietary fiber on PPAR activation in mouse colon. To that end, C57Bl/6 mice were fed either a control diet containing no soluble fiber or the same diet containing 10% inulin for 10 days. Inulin feeding markedly increased concentrations of acetate and propionate in the colon, which reached significance for propionate (Figure 8A). To study the effect of inulin on

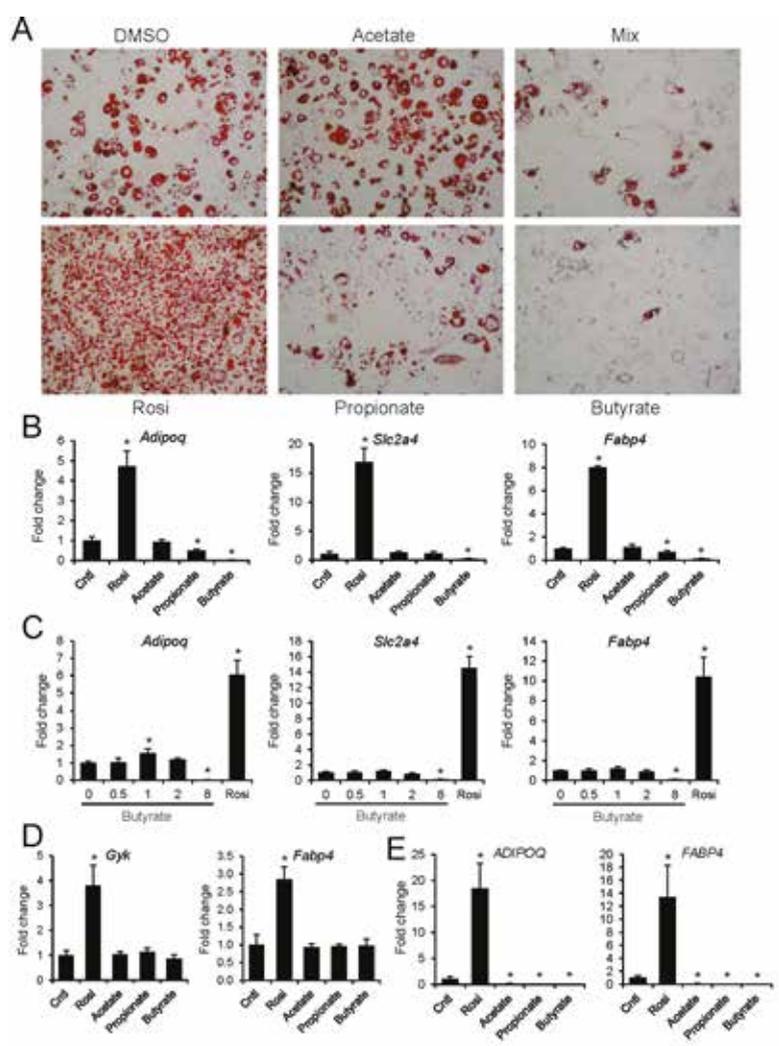


Figure 7 Butyrate and propionate inhibit 3T3-L1 adipogenesis

A) Oil red O staining of 3T3-L1 adipocytes at day 10 treated with SCFA (8 mM) from day 0. Mix contained 2.67 mM of each SCFA. B) Expression of differentiation markers and PPAR γ targets was determined by qPCR at day 4. C) Concentration dependent effect of butyrate on 3T3-L1 differentiation when added at day 0 as determined by expression of differentiation markers at day 4. D) Effect of SCFA (8 mM) and rosiglitazone (1 μ M) on expression of PPAR γ targets in fully differentiated 3T3-L1 adipocytes. Cells were treated for 24 hours. E) Effect of SCFA (8 mM) and rosiglitazone (1 μ M) added on day 1 on human SGBS adipocyte differentiation and expression of adipogenesis marker genes at day 15. Errors bars represent SD. Asterisk indicates significantly different from control according to Student's t-test ($p < 0.05$).

PPAR target gene expression, colons were subjected to microarray analysis followed by gene set enrichment analysis. Remarkably, the most significantly induced gene set was “PPAR targets” (FDR q-value =0, Normalized Enrichment Score =2.44) [54]. It should be noted that PPAR target genes cannot be separated according to PPAR isotype. Representation of the individual expression changes of the most highly ranked genes within “PPAR targets” in a heatmap reveals the pronounced and consistent induction of numerous PPAR targets by inulin (Figure 8B). These data suggest that gut microbial activity and resultant formation of SCFA leads to activation of PPARs.

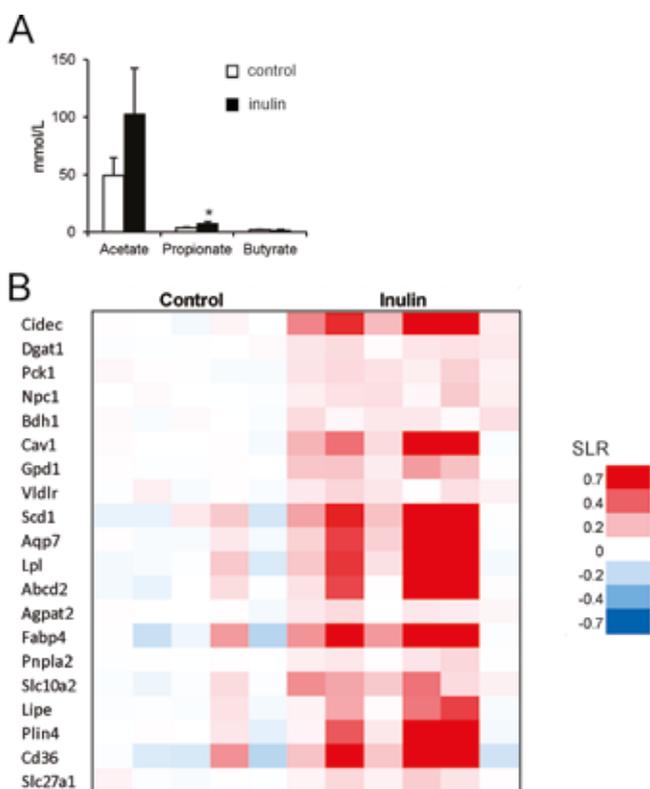


Figure 8 Inulin feeding activates PPAR in colon

Mice were fed a diet enriched with inulin for 10 days. A) Luminal concentration of SCFA in the colon as determined by gas chromatography. Errors bars represent SEM. B) Gene expression changes in colon illustrated by heat map of genes belonging to the most significantly induced gene set “PPAR targets”. SLR: signal log ratio.

Discussion

Here we explored the mechanisms involved in regulation of ANGPTL4 synthesis in human colon. The two major findings are: 1) ANGPTL4 synthesis is highly stimulated by SCFA; 2) SCFA transactivate and bind to PPAR γ , likely by serving as selective PPAR modulators. Overall, the data indicate that SCFA induce *ANGPTL4* mRNA expression and protein secretion in colon cells by activating PPAR γ . Butyrate was the strongest activator followed by propionate, whereas acetate only weakly stimulated PPAR γ and ANGPTL4.

Butyrate is a potent histone deacetylase inhibitor, which likely accounts for most of the observed effects of butyrate in colon adenocarcinoma cells as revealed by microarray. Indeed, a far larger number of genes was regulated by butyrate compared to rosiglitazone, suggesting that PPAR γ activation is quantitatively a relatively minor pathway in gene regulation by SCFA, at least in T84 cells.

Two previous studies have hinted at potential activation of PPAR γ by butyrate [55,56], yet the concept has largely eluded recognition in the field. In contrast, long chain (unsaturated) fatty acids are well-known activators of PPAR γ [57]. They activate PPAR γ at concentrations in the low to medium micromolar range and thus serve as high affinity agonists of PPAR γ . In contrast, concentrations of SCFA needed to activate PPAR γ are in the high micromolar to low millimolar range. Due to the low affinity, the *in vivo* relevance of PPAR γ activation by SCFA is likely insignificant in most human tissues, including adipose tissue. However, the situation is different in the GI-tract and in liver [58]. Indeed, SCFA concentration approaching or even exceeding 100 mM have been reported in human colon and caecum [59]. In mouse intestine we measured total SCFA concentrations of around 40 mM, which would result in substantial activation of PPAR γ . Accordingly, we believe that activation of PPAR γ by SCFA is physiologically only meaningful in human caecum and colon, and perhaps in liver.

Interestingly, medium chain fatty acids (MCFA, C8-C10) were recently shown to act as modulators of PPAR γ [31]. Similar to the data reported here for SCFA, the MCFA decanoic acid bound and (trans)activated PPAR γ , and contrary to synthetic PPAR γ agonists and LCFA, decanoic acid inhibited adipogenesis. Furthermore, it was shown that the hydrocarbon tail of decanoic acid occupies a completely different pocket compared to the tail of LCFA or rosiglitazone. Remarkably, even though no specific orientational or positional restraints were used to guide the binding of butyrate in the large PPAR γ binding site, the most favorable solution to the docking was very similar to that of decanoic acid. In general, the best solutions of the docking protocol displayed high quality interactions with the PPAR γ receptor, with a slightly different orientation of the butyrate carboxylic acid group in the binding site compared to decanoic acid. The model shows that butyrate is stabilized in the binding site by interactions with protein side-chains. Since the buried surface area and interaction energy with the receptor are less than for decanoic acid, the affinity of the complex with butyrate is predicted to be weaker than with decanoic acid.

PPAR γ has an antineoplastic effect in many different tumor types, yet its role in colorectal tumors remains controversial [60]. In contrast, the anti-inflammatory effect of PPAR γ in the colon is well recognized [61]. PPAR γ ligands were shown to suppress inflammatory gene expression in colonic cell lines by suppressing NF- κ B and reduce inflammation in a mouse model of Inflammatory Bowel Disease [62,63]. In addition, PPAR γ in colonic epithelial cells was shown to protect against experimental inflammatory bowel disease [49]. Similarly, SCFA, especially butyrate, seem to have broad anti-inflammatory properties by altering immune cell migration, adhesion, and cytokine expression, and by affecting cell proliferation and apoptosis [64]. Accordingly, it can be hypothesized that the anti-inflammatory properties of SCFA in the colon are at least partially conveyed by PPAR γ [65].

Based on data presented here, microbiota may be able to influence ANGPTL4 production via production of SCFA and subsequent activation of PPAR γ . Backhed reported that colonization of the gut of germ free mice with microbiota reduces *ANGPTL4* expression in mouse intestine [2]. In as much as SCFA stimulate *ANGPTL4* expression, the suppressive effect of colonization on ANGPTL4 must be mediated by a mechanism other than SCFA. It has been suggested that alterations in intestinal *ANGPTL4* expression may influence adipose LPL activity and thereby impact adipose mass [2]. SCFA may thus inhibit fat storage by stimulating release of ANGPTL4. Whether ANGPTL4 also has a functional role in the intestine is unclear. Since the intestine does not express LPL, the local role of ANGPTL4 in intestine must extend beyond LPL inhibition. LPL-inhibition is conferred exclusively by the N-terminal domain of ANGPTL4, whereas the C-terminal fragment of ANGPTL4 acts as a ligand for integrins to alter cellular signaling [66-68].

In conclusion, we show that SCFA potently stimulate ANGPTL4 production in human colon cell lines via PPAR γ . Our data point to activation of PPARs as a novel mechanism of gene regulation by SCFA in the colon, in addition to other mechanisms of action of SCFA.

Acknowledgements

This study was supported by the Netherlands Nutrigenomics Centre, the Netherlands Organization for Scientific Research (NWO) (VICI grant 700.56.442 to AB, visitor travel grant 040.11.299 to JG, TOP grant 40-00812-98-08030 to SK), and the Netherlands Consortium for Systems Biology.

References

1. Sekirov I, Russell SL, Antunes LC, Finlay BB (2010) Gut microbiota in health and disease. *Physiological Reviews* 90: 859-904.
2. Backhed F, Ding H, Wang T, Hooper LV, Koh GY, et al. (2004) The gut microbiota as an environmental factor that regulates fat storage. *Proc Natl Acad Sci U S A* 101: 15718-15723.
3. Backhed F, Manchester JK, Semenkovich CF, Gordon JI (2007) Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. *Proc Natl Acad Sci U S A* 104: 979-984.
4. Lichtenstein L, Kersten S (2010) Modulation of plasma TG lipolysis by Angiopoietin-like proteins and GPIHBP1. *Biochim Biophys Acta* 1801: 415-420.
5. Yoshida K, Shimizugawa T, Ono M, Furukawa H (2002) Angiopoietin-like protein 4 is a potent hyperlipidemia-inducing factor in mice and inhibitor of lipoprotein lipase. *J Lipid Res* JID - 0376606 43: 1770-1772.
6. Ge H, Yang G, Yu X, Pourbahrami T, Li C (2004) Oligomerization state-dependent hyperlipidemic effect of angiopoietin-like protein 4. *J Lipid Res* 45: 2071-2079.
7. Koster A, Chao YB, Mosior M, Ford A, Gonzalez-DeWhitt PA, et al. (2005) Transgenic angiopoietin-like (angptl)4 overexpression and targeted disruption of angptl4 and angptl3: regulation of triglyceride metabolism. *Endocrinology* 146: 4943-4950.
8. Mandard S, Zandbergen F, van Straten E, Wahli W, Kuipers F, et al. (2006) The Fasting-induced Adipose Factor/Angiopoietin-like Protein 4 Is Physically Associated with Lipoproteins and Governs Plasma Lipid Levels and Adiposity. *J Biol Chem* 281: 934-944.
9. Xu A, Lam MC, Chan KW, Wang Y, Zhang J, et al. (2005) Angiopoietin-like protein 4 decreases blood glucose and improves glucose tolerance but induces hyperlipidemia and hepatic steatosis in mice. *Proc Natl Acad Sci U S A* 102: 6086-6091.
10. Yu X, Burgess SC, Ge H, Wong KK, Nassef RH, et al. (2005) Inhibition of cardiac lipoprotein utilization by transgenic overexpression of Angptl4 in the heart. *Proc Natl Acad Sci U S A* 102: 1767-1772.
11. Georgiadi A, Lichtenstein L, Degenhardt T, Boekschoten MV, van Bilsen M, et al. (2010) Induction of cardiac Angptl4 by dietary fatty acids is mediated by peroxisome proliferator-activated receptor beta/delta and protects against fatty acid-induced oxidative stress. *Circ Res* 106: 1712-1721.
12. Lichtenstein L, Mattijssen F, de Wit NJ, Georgiadi A, Hooiveld GJ, et al. (2010) Angptl4 protects against severe proinflammatory effects of saturated fat by inhibiting fatty acid uptake into mesenteric lymph node macrophages. *Cell Metab* 12: 580-592.
13. Kersten S, Mandard S, Tan NS, Escher P, Metzger D, et al. (2000) Characterization of the fasting-induced adipose factor FIAF, a novel peroxisome proliferator-activated receptor target gene. *J Biol Chem* 275: 28488-28493.
14. Yoon JC, Chickering TW, Rosen ED, Dussault B, Qin Y, et al. (2000) Peroxisome proliferator-activated receptor gamma target gene encoding a novel angiopoietin-related protein associated with adipose differentiation. *Mol Cell Biol* 20: 5343-5349.
15. Bunger M, van den Bosch HM, van der Meijde J, Kersten S, Hooiveld GJ, et al. (2007) Genome-wide analysis of PPARalpha activation in murine small intestine. *Physiol Genomics* 30: 192-204.
16. Staiger H, Haas C, Machann J, Werner R, Weisser M, et al. (2009) Muscle-derived angiopoietin-like protein 4 is induced by fatty acids via peroxisome proliferator-activated receptor (PPAR)-delta and is of metabolic relevance in humans. *Diabetes* 58: 579-589.
17. Kersten S, Lichtenstein L, Steenbergen E, Mudde K, Hendriks HF, et al. (2009) Caloric restriction and exercise increase plasma ANGPTL4 levels in humans via elevated free fatty acids. *Arterioscler Thromb Vasc Biol* 29: 969-974.
18. Modica S, Gofflot F, Murzilli S, D'Orazio A, Salvatore L, et al. (2010) The intestinal nuclear receptor signature with epithelial localization patterns and expression modulation in tumors. *Gastroenterology* 138: 636-648, 648 e631-612.
19. Bookout AL, Jeong Y, Downes M, Yu RT, Evans RM, et al. (2006) Anatomical profiling of nuclear receptor expression reveals a hierarchical transcriptional network. *Cell* 126: 789-799.
20. Aronsson L, Huang Y, Parini P, Korach-Andre M, Hakansson J, et al. (2010) Decreased fat storage by *Lactobacillus paracasei* is associated with increased levels of angiopoietin-like 4 protein (ANGPTL4). *PLoS One* 5.

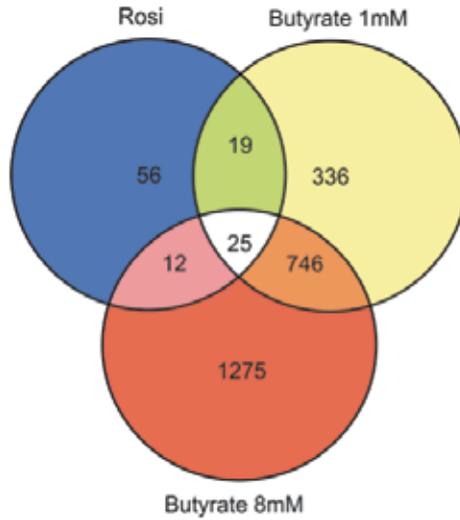
21. Siersbaek MS, Loft A, Aagaard MM, Nielsen R, Schmidt SF, et al. (2012) Genome-wide profiling of peroxisome proliferator-activated receptor gamma in primary epididymal, inguinal, and brown adipocytes reveals depot-selective binding correlated with gene expression. *Mol Cell Biol* 32: 3452-3463.
22. Gijbbers L, Man HY, Kloet SK, de Haan LH, Keijer J, et al. (2011) Stable reporter cell lines for peroxisome proliferator-activated receptor gamma (PPARgamma)-mediated modulation of gene expression. *Anal Biochem* 414: 77-83.
23. Sonneveld E, Riteco JA, Jansen HJ, Pieterse B, Brouwer A, et al. (2006) Comparison of in vitro and in vivo screening models for androgenic and estrogenic activities. *Toxicol Sci* 89: 173-187.
24. Sonneveld E, Jansen HJ, Riteco JA, Brouwer A, van der Burg B (2005) Development of androgen- and estrogen-responsive bioassays, members of a panel of human cell line-based highly selective steroid-responsive bioassays. *Toxicol Sci* 83: 136-148.
25. Koppen A, Houtman R, Pijnenburg D, Jeninga EH, Ruijtenbeek R, et al. (2009) Nuclear receptor-coregulator interaction profiling identifies TRIP3 as a novel peroxisome proliferator-activated receptor gamma cofactor. *Molecular & cellular proteomics : MCP* 8: 2212-2226.
26. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5: R80.
27. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, et al. (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4: 249-264.
28. Dai M, Wang P, Boyd AD, Kostov G, Athey B, et al. (2005) Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res* 33: e175.
29. Scarsi M, Podvinec M, Roth A, Hug H, Kersten S, et al. (2007) Sulfonylureas and glinides exhibit peroxisome proliferator-activated receptor gamma activity: a combined virtual screening and biological assay approach. *Mol Pharmacol* 71: 398-406.
30. Dominguez C, Boelens R, Bonvin AM (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* 125: 1731-1737.
31. Malapaka RR, Khoo S, Zhang J, Choi JH, Zhou XE, et al. (2012) Identification and mechanism of 10-carbon fatty acid as modulating ligand of peroxisome proliferator-activated receptors. *J Biol Chem* 287: 183-195.
32. de Vries SJ, van Dijk M, Bonvin AM (2010) The HADDOCK web server for data-driven biomolecular docking. *Nat Protoc* 5: 883-897.
33. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, et al. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* 54: 905-921.
34. Jorgensen WL, Tirado-Rives J (1988) The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J Am Chem Soc* 110: 1657-1666.
35. Macfarlane GT, Gibson GR, Cummings JH (1992) Comparison of fermentation reactions in different regions of the human colon. *The Journal of applied bacteriology* 72: 57-64.
36. Boffa LC, Vidali G, Mann RS, Allfrey VG (1978) Suppression of histone deacetylation in vivo and in vitro by sodium butyrate. *J Biol Chem* 253: 3364-3366.
37. Waldecker M, Kautenburger T, Daumann H, Busch C, Schrenk D (2008) Inhibition of histone-deacetylase activity by short-chain fatty acids and some polyphenol metabolites formed in the colon. *J Nutr Biochem* 19: 587-593.
38. Harmon GS, Dumlao DS, Ng DT, Barrett KE, Dennis EA, et al. (2010) Pharmacological correction of a defect in PPAR-gamma signaling ameliorates disease severity in Cfrt-deficient mice. *Nat Med* 16: 313-318.
39. Han X, Benight N, Osuntokun B, Loesch K, Frank SJ, et al. (2007) Tumour necrosis factor alpha blockade induces an anti-inflammatory growth hormone signalling pathway in experimental colitis. *Gut* 56: 73-81.
40. Lee SK, Kim YW, Chi SG, Joo YS, Kim HJ (2009) The effect of *Saccharomyces boulardii* on human colon cells and inflammation in rats with trinitrobenzene sulfonic acid-induced colitis. *Dig Dis Sci* 54: 255-263.
41. Narravula S, Colgan SP (2001) Hypoxia-inducible factor 1-mediated inhibition of peroxisome proliferator-activated receptor alpha expression during hypoxia. *J Immunol* 166: 7543-7548.
42. Shimada T, Kojima K, Yoshiura K, Hiraishi H, Terano A (2002) Characteristics of the peroxisome proliferator activated receptor gamma (PPARgamma) ligand induced apoptosis in colon cancer cells. *Gut* 50: 658-664.
43. Wang JB, Qi LL, Zheng SD, Wang HZ, Wu TX (2009) Curcumin suppresses PPARdelta expression and related genes in HT-29 cells. *World J Gastroenterol* 15: 1346-1352.

44. Kaddatz K, Adhikary T, Finkernagel F, Meissner W, Muller-Brusselbach S, et al. (2010) Transcriptional profiling identifies functional interactions of TGF beta and PPAR beta/delta signaling: synergistic induction of ANGPTL4 transcription. *J Biol Chem* 285: 29469-29479.
45. Mandard S, Zandbergen F, Tan NS, Escher P, Patsouris D, et al. (2004) The direct peroxisome proliferator-activated receptor target fasting-induced adipose factor (FIAF/PGAR/ANGPTL4) is present in blood plasma as a truncated protein that is increased by fenofibrate treatment. *J Biol Chem* 279: 34411-34420.
46. Camirand A, Marie V, Rabelo R, Silva JE (1998) Thiazolidinediones stimulate uncoupling protein-2 expression in cell lines representing white and brown adipose tissues and skeletal muscle. *Endocrinology* 139: 428-431.
47. Kronke G, Kadl A, Ikonomu E, Bluml S, Furnkranz A, et al. (2007) Expression of heme oxygenase-1 in human vascular cells is regulated by peroxisome proliferator-activated receptors. *Arterioscler Thromb Vasc Biol* 27: 1276-1282.
48. Wu P, Inskip K, Bowker-Kinley MM, Popov KM, Harris RA (1999) Mechanism responsible for inactivation of skeletal muscle pyruvate dehydrogenase complex in starvation and diabetes. *Diabetes* 48: 1593-1599.
49. Adachi M, Kurotani R, Morimura K, Shah Y, Sanford M, et al. (2006) Peroxisome proliferator activated receptor gamma in colonic epithelial cells protects against experimental inflammatory bowel disease. *Gut* 55: 1104-1113.
50. Lytle C, Tod TJ, Vo KT, Lee JW, Atkinson RD, et al. (2005) The peroxisome proliferator-activated receptor gamma ligand rosiglitazone delays the onset of inflammatory bowel disease in mice with interleukin 10 deficiency. *Inflamm Bowel Dis* 11: 231-243.
51. Higgins LS, Depaoli AM (2010) Selective peroxisome proliferator-activated receptor gamma (PPARgamma) modulation as a strategy for safer therapeutic PPARgamma activation. *Am J Clin Nutr* 91: 267S-272S.
52. Guan HP, Ishizuka T, Chui PC, Lehrke M, Lazar MA (2005) Corepressors selectively control the transcriptional activity of PPARgamma in adipocytes. *Genes Dev* 19: 453-461.
53. Yu C, Markan K, Temple KA, Deplewski D, Brady MJ, et al. (2005) The nuclear receptor corepressors NCoR and SMRT decrease peroxisome proliferator-activated receptor gamma transcriptional activity and repress 3T3-L1 adipogenesis. *J Biol Chem* 280: 13600-13605.
54. Rakhshandehroo M, Knoch B, Muller M, Kersten S (2010) Peroxisome proliferator-activated receptor alpha target genes. *PPAR Res* 2010.
55. Kinoshita M, Suzuki Y, Saito Y (2002) Butyrate reduces colonic paracellular permeability by enhancing PPARgamma activation. *Biochem Biophys Res Commun* 293: 827-831.
56. Schwab M, Reynders V, Loitsch S, Steinhilber D, Stein J, et al. (2007) Involvement of different nuclear hormone receptors in butyrate-mediated inhibition of inducible NF kappa B signalling. *Mol Immunol* 44: 3625-3632.
57. Georgiadi A, Kersten S (2012) Mechanisms of gene regulation by Fatty acids. *Adv Nutr* 3: 127-134.
58. Bugaut M, Bentejac M (1993) Biological effects of short-chain fatty acids in nonruminant mammals. *Annu Rev Nutr* 13: 217-241.
59. Cummings JH, Pomare EW, Branch WJ, Naylor CP, Macfarlane GT (1987) Short chain fatty acids in human large intestine, portal, hepatic and venous blood. *Gut* 28: 1221-1227.
60. Dai Y, Wang WH (2010) Peroxisome proliferator-activated receptor gamma and colorectal cancer. *World J Gastrointest Oncol* 2: 159-164.
61. Dubuquoy L, Rousseaux C, Thuru X, Peyrin-Biroulet L, Romano O, et al. (2006) PPARgamma as a new therapeutic target in inflammatory bowel diseases. *Gut* 55: 1341-1349.
62. Kelly D, Campbell JJ, King TP, Grant G, Jansson EA, et al. (2004) Commensal anaerobic gut bacteria attenuate inflammation by regulating nuclear-cytoplasmic shuttling of PPAR-gamma and RelA. *Nat Immunol* 5: 104-112.
63. Su CG, Wen X, Bailey ST, Jiang W, Rangwala SM, et al. (1999) A novel therapy for colitis utilizing PPAR-gamma ligands to inhibit the epithelial inflammatory response. *J Clin Invest* 104: 383-389.
64. Meijer K, de Vos P, Priebe MG (2010) Butyrate and other short-chain fatty acids as modulators of immunity: what relevance for health? *Curr Opin Clin Nutr Metab Care* 13: 715-721.
65. Hamer HM, Jonkers D, Venema K, Vanhoutvin S, Troost FJ, et al. (2008) Review article: the role of butyrate on colonic function. *Aliment Pharmacol Ther* 27: 104-119.
66. Goh YY, Pal M, Chong HC, Zhu P, Tan MJ, et al. (2010) Angiopoietin-like 4 interacts with integrins beta1 and beta5 to modulate keratinocyte migration. *Am J Pathol* 177: 2791-2803.
67. Goh YY, Pal M, Chong HC, Zhu P, Tan MJ, et al. (2010) Angiopoietin-like 4 interacts with matrix proteins to modulate wound healing. *J Biol Chem* 285: 32999-33009.



68. Sukonina V, Lookene A, Olivecrona T, Olivecrona G (2006) Angiotensin-like protein 4 converts lipoprotein lipase to inactive monomers and modulates lipase activity in adipose tissue. *Proc Natl Acad Sci U S A* 103: 17450-17455.
69. Xu HE, Stanley TB, Montana VG, Lambert MH, Shearer BG, et al. (2002) Structural basis for antagonist-mediated recruitment of nuclear co-repressors by PPARalpha. *Nature* 415: 813-817.

Supplemental Information



Supplemental Figure 1

Venn diagram showing overlap in gene induction by butyrate (1 mM or 8 mM) and rosiglitazone (1 μ M) in T84 cells based on Affymetrix microarray analysis. Only genes with signal intensity >20 and number of probesets >10 were included. Threshold for regulation was set at fold change of 1.8.



Chapter 6

Role of epithelial peroxisome proliferator-activated receptor (PPAR) γ in colonic response to fermentation of inulin in mice

Katja Lange, Floor Hugenholtz*, Noortje IJssennagger*, Marcela Doktorova, Joep van Bortel, Jacques Vervoort, Saskia van Mil, Michiel Kleerebezem, Hauke Smidt, Hans Jonker, Michael Müller, Guido J.E.J. Hooiveld

*These authors contributed equally to this work

In preparation

Abstract

Dietary fibers are known to promote gastrointestinal homeostasis. Dietary fibers are fermented in the large intestine by microbiota to produce short-chain fatty acids (SCFA). SCFA, in particular butyrate, transactivate peroxisome proliferator-activated receptor (PPAR) γ . Studies with mice deficient in intestinal PPAR γ suggested an important role for PPAR γ in gastrointestinal homeostasis. Target genes of PPAR γ were activated by dietary fibers that are fermented.

Here we addressed the question which role PPAR γ in the intestinal epithelium played in the fermentation of dietary fiber in mice. Mice lacking epithelial *Ppar γ* (KO) in intestine (via *cre-vilin*) and wild type (WT) mice were fed the soluble dietary fiber inulin for 10 days. Subsequently, whole-genome gene expression of mucosal cells was analyzed by microarray analysis. The colonic content was analyzed for microbiota composition using mouse intestinal tract chip (MITChip) and metabolite level were measured by ¹H-NMR spectroscopy. In addition, colonic crypt cells from both KO and WT mice were grown to organoids and subsequently treated with 1mM butyrate for 24hrs, followed by whole-genome gene expression analysis to identify butyrate mediated *Ppar γ* -dependently regulated genes.

We found that diet more than the genotype status of the mice coincides with variation in metabolite levels, microbiota composition and gene expression in colon. We identified *Ppar γ* -dependent genes regulated by inulin indicating that *Ppar γ* partially governs colonic response to fermentation. These genes clustered in biological processes related to oxidative stress response, catabolic metabolism, DNA repair, cell cycle, immune related processes, insulin signaling and gene transcription. In addition, *Ppar γ* -dependent gene expression response to butyrate treated organoids was related to energy-generating processes, *Nrf2* targets, amino acid and glucose metabolism. In conclusion, epithelial *Ppar γ* partially mediates gene expression response to inulin and butyrate in colon.



Background

Dietary fibers are known to promote gastrointestinal homeostasis [1-4]. Dietary fibers such as inulin are fermented by the intestinal microbiota. Fermentation results in the production of microbial metabolites among which short-chain fatty acids (SCFA) are the most abundant ones. Among the SCFA, acetate, propionate and butyrate are produced in high concentrations in the murine large intestine [5]. Beneficial effects of dietary fibers can in part be attributed to SCFA. Butyrate is thought to be the primary energy source for colonocytes [6]. Although propionate and acetate are mainly transferred and taken up by other organs than the large intestine, it has been demonstrated that they also impact colonic gene expression in mice [chapter 2]. Previously, it was found that dietary fibers that increased SCFA concentrations in the colonic luminal content of mice commonly increased the expression of Ppar target genes in murine colon [chapter 3]. Among the different dietary fibers, in particular inulin showed strongest activation of Ppar γ as suggested by bioinformatics analyses. Additionally, it was reported that butyrate and to lesser extent propionate specifically transactivate the isoform PPAR γ , as was shown with transactivation assays [7] using physiological concentrations of SCFA [chapter 3]. Hence, Ppar γ was identified as potential regulator in the fiber fermentation process mediating effects of dietary fiber on gastrointestinal homeostasis. Only a few studies have been published that studied the role of PPAR γ in the intestine. It is well known that PPAR γ signaling is crucial for adipogenesis, insulin sensitivity and immunity [8]. Intestine-specific knock-out (KO) models of *Ppar γ* have been described in literature and showed that Ppar γ is important for maintaining intestinal homeostasis by affecting tumorigenesis [9] and mucosal immune response [10]. The role of PPAR γ in the process of fermentation of dietary fiber is not known.

Therefore, we used intestinal epithelial-specific *Ppar γ* KO mice, which were fed with inulin for 10 days, followed by analyses of intestinal physiology including microbiota composition, SCFA, and whole-genome gene expression of mucosal cell scrapings to elucidate on the role Ppar γ in fermentation process. To study Ppar γ -dependent effects of butyrate on gene expression, colonic crypt cells derived from *Ppar γ* KO or wild- type mice were grown *ex vivo* to form organoids [11] and treated with SCFA.

Material and Methods

Mice

Male C57BL/6J mice were housed in a light- and temperature-controlled animal facility of the University Medical Centre Groningen (light on from 7AM to 7PM, 21 °C) with free access to food and water. Intestine-specific *Ppar γ* knock out (Ppar $\gamma^{\Delta EC}$) mice were generated from mice which had loxP-flanked exon 1 and 2 of the *Ppar γ* gene (developed in Evans Laboratory and described in [12]). These mice were backcrossed with transgenic

mice expressing Cre recombinase under control of the villin promoter. Mice (n=8 per genotype) were fed either low-fat diet (LFD) or inulin (IN) supplemented diets (10% of corn starch was replaced by IN) for 10 days (diet composition are given in **chapter 3**). The experiment was approved by the Ethics Committee for Care and Use of Laboratory Animals of the University of Groningen. On the day of sections, mice were fasted for 2hrs. Two hours later mice were anaesthetized with isoflurane, and the colon was excised. The adhering fat around the colon was carefully removed, and the colon was cut open longitudinally. The luminal content was sampled and the tissue was rinsed with ice-cold phosphate buffered saline. Subsequently, the epithelial lining of the colon was scraped off. Luminal content and scrapings were collected in tubes, which were immediately snap frozen in liquid nitrogen and stored at -80°C.

Organoids

From 3 WT and 3 Ppar $\gamma^{\Delta IEC}$ mice colon was isolated, washed with cold PBS, epithelial layer was gently scraped, the remaining tissue with intact crypts was washed several times with cold PBS and subsequently incubated for 1hr in 5mM EDTA in PBS to isolate crypt cells as previously described [11]. The washed and filtered crypt cells were embedded in Matrigel and seeded in a 24-well plates containing advanced DMEM/F12 supplemented with Penicillin/Streptomycin, 10 mM Hepes, Glutamax, Wnt, nAcetyl-cysteine, growth factors (Noggin, R-spondin, mEGF), B27, kinase inhibitor, Nicotinamide, A83 and p38.

Butyrate treatment. Organoids were treated with butyrate (1mM) for 24hrs. PBS served as control. After 24 hours organoids were harvested and stored in Trizol at -80C until RNA isolation.

MITChip

Colonic content of 20 mice (n=5 per group) was used for microbiota composition analysis by MITChip following protocols and data analyses as described before [**chapter 4**].

¹H-NMR spectroscopy

Colonic and cecal content samples were collected, homogenized in 100mM phosphoric buffer solution (pH 8) and kept at -20 C. Samples were then diluted in phosphoric buffer with 1mM maleic acid as standard. Subsequently, 200 μ L were transferred to a 3 mm NMR tube (Bruker match system). Samples were measured at 310 K (calibrated temperature) in an Avance III NMR spectrometer operated at 600.13 MHz as described in [13].

Microarray analysis and data processing

Colonic scrapings (n=6 per diet group) were subjected to genome-wide expression profiling as described before [**chapter 3**]. After quality control 1 array from the control group had to be excluded from further analysis due to insufficient quality. The dataset was filtered to only include probe sets that were active (i.e. expressed) in at least 3 samples



using the universal expression code (UPC) approach (UPC score > 0.50). This resulted in the inclusion of 9,329 of the 21,187 probe sets. For organoid tissues, 9,354 of the 21,187 probe sets were included. Differentially expressed probe sets were identified by linear models and an intensity-based moderated t-test [14,15]. Probe sets that satisfied the criterion of $P < 0.01$ were considered to be significantly regulated.

Statistical analysis

Statistical differences were determined with two-way ANOVA using GraphPad Prism version 5.04 Windows, GraphPad Software, San Diego California USA. Differences were considered significant at * $P \leq 0.05$ or ** $P \leq 0.01$. Data are presented as mean \pm SEM.

Results

I. The role of intestinal, epithelial *Ppar γ* on effects of inulin

Physiological parameter

Feeding inulin and a LF control diet to WT and *Ppar γ ^{Δ IEC}* mice did not lead to significant differences in cumulative food intake or body weight after the 10 days feeding period (**Figure 1**). The total cecum showed significantly higher weight in mice fed inulin for both genotypes, whereas for colon no significant differences between the groups were observed (**S Figure 1**).

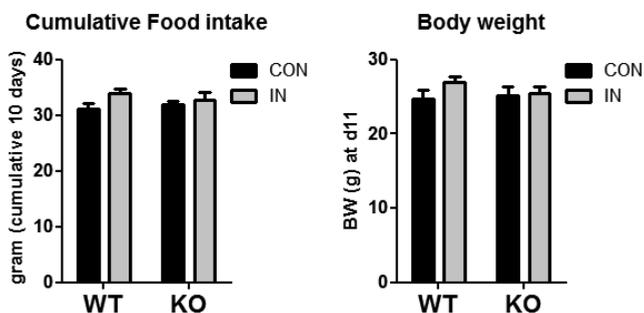


Figure 1

Food intake measured over 10 days feeding control diet (CON) (low fat diet) or inulin supplemented diet (IN) in wild type (WT) or KO (*cre-vil Ppar γ KO*) mice and body weight at the end of the feeding period.

Microbiota composition changes in colonic content

To analyze the effect of inulin and intestinal *Pparγ* on microbiota composition, a MITChip analysis was performed. Microbiota composition in mice fed the same diet was more similar than mice with different genotype status (**Figure 2**). The four different experimental groups explained 22.3% of the total variation in composition. In addition, samples of *Pparγ*^{ΔIEC} mice fed with IN showed a high variation. Although we can conclude a predominant role of diet on microbiota composition rather than epithelial *Pparγ*, there are also bacterial groups modulated in *Pparγ*^{ΔIEC} mice indicating a role of *Pparγ* on composition.

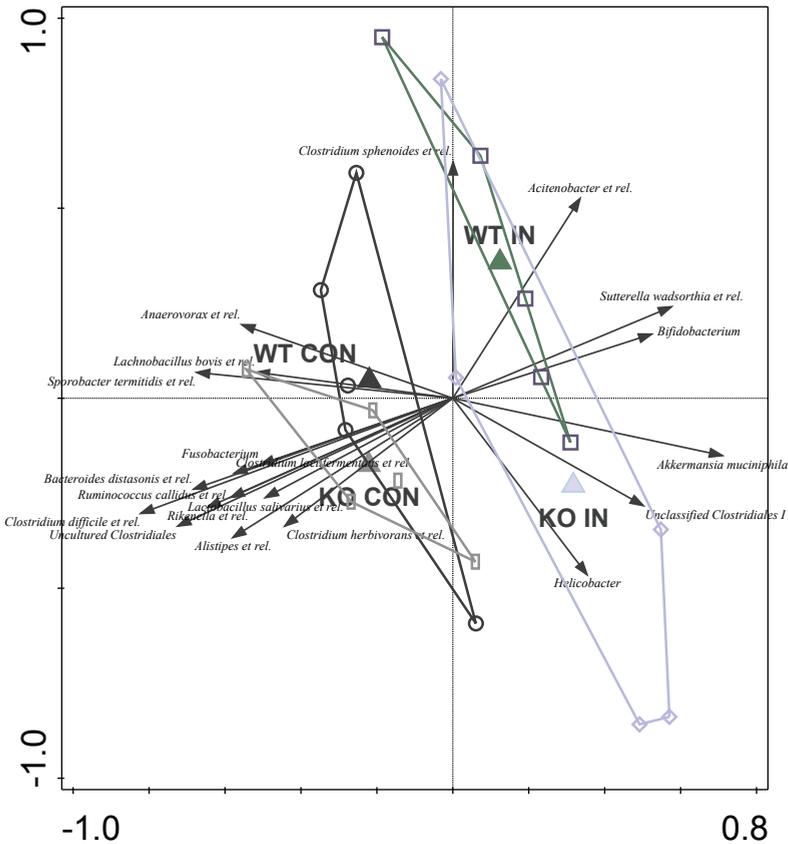


Figure 2

Redundancy analysis of colon microbiota from wildtype (WT) or *Pparγ* KO (KO) mice on a control or inulin diet. The variables were the different groups - WT on control diet (WT control); WT on inulin diet (WT Inulin); *Pparγ* KO on control diet and *Pparγ* KO on the inulin diet. These variables together accounted for 22.3% of the total variation, of which in this plot 92.2% is shown.



Metabolomics changes in colonic content

Metabolomics was used to determine compounds in luminal content of colon. After applying Partial Least Square –Discriminant Analysis (PLS-DA) on the 214 binned compounds measured by $^1\text{H-NMR}$, we observed clustering of samples from mice receiving inulin vs. control diets on the second and third component, independent of the genotype, not however on the first component (**Figure 3A**). SCFA levels were determined since these compounds are expected to be formed by fermentation and are known to activate PPAR γ . Diet had a significant effect on acetate levels, whereas there was no significant genotype effect. Acetate and propionate levels were higher in mice fed IN compared to control diet fed mice for both genotypes (**Figure 3B**). For butyrate levels, however, there was a significant effect of interaction between diet and genotype. While levels were higher in Ppar $\gamma^{\Delta\text{IEC}}$ with IN fed mice, they were lower for WT mice. In addition, we noticed high variation between Ppar $\gamma^{\Delta\text{IEC}}$ mice fed IN. Taken together, our data suggest an effect of intestinal Ppar γ on specific luminal metabolite peak integrals, in particular for butyrate, although overall the diet effect seemed to be larger than the genotype effect.

Ppar γ dependent gene expression changes induced by inulin

The 9,329 filtered genes were subjected to PLS-DA to compare gene expression profiles between mice. Overall, gene expression profiles were more comparable between mice fed the same diet than between mice from the same genotype indicating that the diet effect is more pronounced than the genotype effect (**Figure 4A**). To analyze the role of genotype effect, genes significantly regulated between IN and control fed WT and Ppar $\gamma^{\Delta\text{IEC}}$ mice, respectively were determined ($P < 0.01$). In IN fed mice 449 genes were significantly regulated, while in Ppar $\gamma^{\Delta\text{IEC}}$ mice 171 genes were regulated suggesting that abolishing intestinal Ppar γ dampened part of the effect of IN on the colonic transcriptome (**Figure 4B**). To determine IN induced changes mediated via Ppar γ , Venn plots were used to calculate the number of genes only regulated in WT mice by IN and not in Ppar $\gamma^{\Delta\text{IEC}}$ mice. From the 449 genes regulated by IN under WT conditions, 397 genes were regulated by IN only in the WT mice, whereas 52 genes were also regulated in the Ppar $\gamma^{\Delta\text{IEC}}$ mice and 119 genes were only regulated in Ppar $\gamma^{\Delta\text{IEC}}$ mice (**Figure 4C**). The large number of genes (~88%) regulated only in Ppar γ WT mice indicated that Ppar γ plays an important role in mediating IN induced effects on colonic transcriptome.

In addition, there were two sets of genes regulated with IN in WT mice but not in KO mice, suggesting Ppar γ dependency. (1) Genes regulated by IN only in WT that were also regulated between WT and Ppar $\gamma^{\Delta\text{IEC}}$ mice fed the control diet (43 genes) suggesting additional mechanism. (2) The majority of genes that were regulated by IN only in WT mice were not regulated between WT and Ppar $\gamma^{\Delta\text{IEC}}$ fed the control diet (354 genes) suggesting direct Ppar γ -dependent regulation. The highest regulated genes with IN vs. control diet from (1) were *Reg3g* (FC 4.09), *Reg3b* (FC 3.49) and from (2) *Ly6g6c* ($P < 0.01$, FC 2.69) and *Wfdc18* ($P < 0.01$, FC 2.13) (**Figure 4D**).

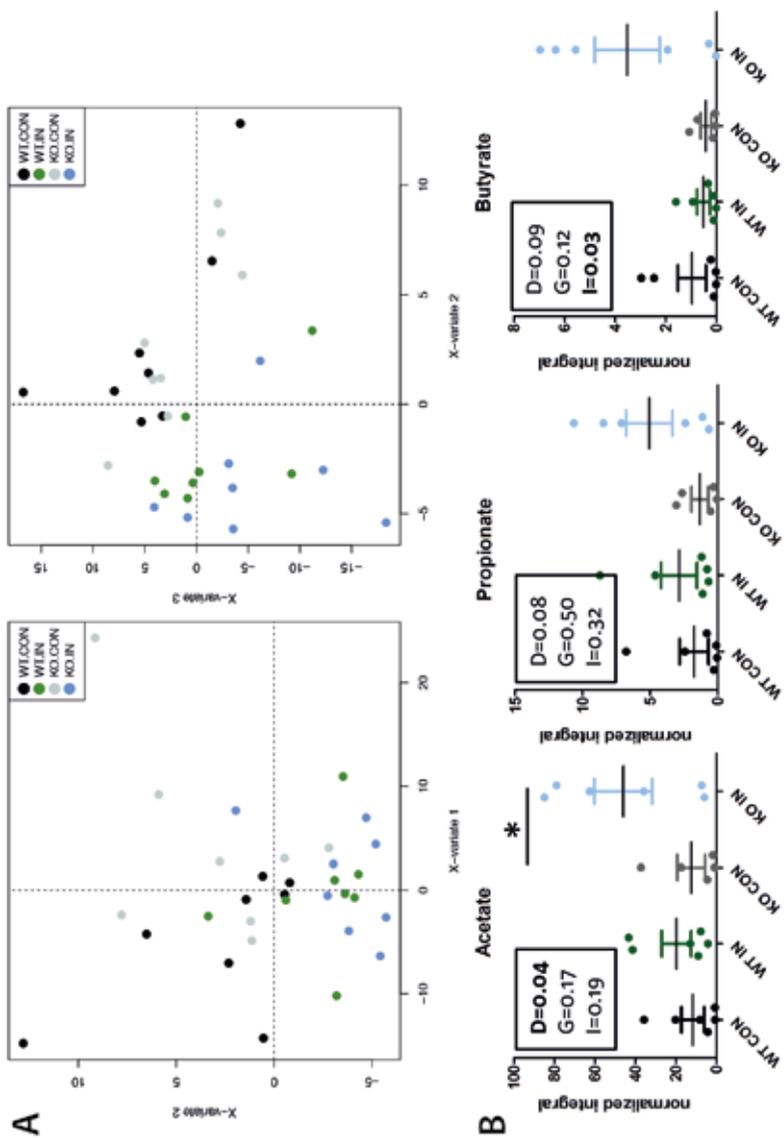


Figure 3

Metabolite level in content of colon measured with $^1\text{H-NMR}$; Partial least square analysis showing variation of mice based on integrals of components (A); Average with standard error of integrals of SCFA for (B), two-way ANOVA for diet effect (D), genotype effect (G), or interaction (I),* indicates significant difference between two groups with $P < 0.05$; WT=wild type, KO= cre-vil *Ppar γ* KO, IN=inulin, CON=control diet

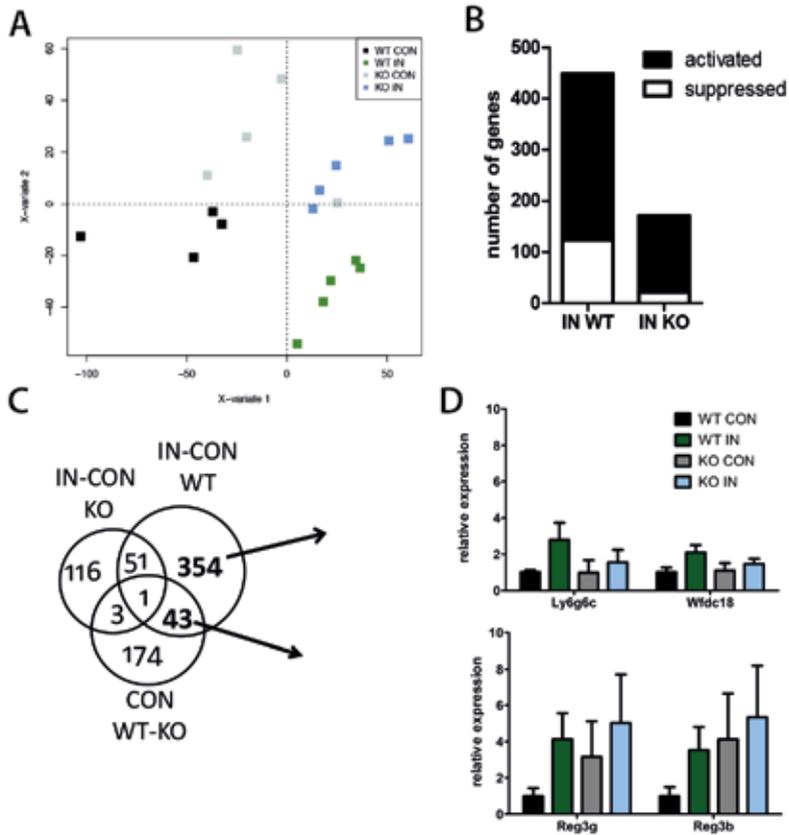


Figure 4

Gene expression profile in colonic tissue scrapings measured with microarray; A) Partial least square analysis showing variation of mice based on gene expression; B) Number of genes activated and suppressed in inulin fed mice compared to control diet fed WT (wild type) and KO (cre-vil *Ppary* KO) mice C) Venn plots comparing number of genes significantly regulated ($P < 0.01$) D) Genes regulated between IN and CON fed mice and not regulated between WT and KO mice (upper) and genes also regulated between WT and KO mice

Gene sets regulated by inulin

Using gene set enrichment analysis (GSEA), the potential role of *Ppary* in mediating effects of IN on intestinal epithelial cells was further explored. Gene sets that were significantly regulated ($P < 0.001$, $FDR < 0.05$) in IN fed WT mice but were not regulated in *Ppary*^{ΔIEC} mice were selected and presented in an enrichment map (**Figure 5**). Among the biological

processes that were enriched with Ppar γ -dependent, activated genes were catabolic processes (TCA cycle, electron transport chain, redox reactions), phase I and II metabolism and Nrf2 target genes, DNA repair and cell cycle (**Figure 5A**). Processes enriched with Ppar γ dependent suppressed genes encompassed immune related processes (interferon signaling, T cell signaling, cytokine metabolic process, antigen processing and presentation), insulin signaling and gene transcription (**Figure 5B**). The differentially regulated gene sets indicate a role for epithelial Ppar γ in the activation of metabolic processes and suppression of immune-related processes mainly.

II. Ppar γ dependent SCFA effects on colonic gene expression

It has been reported before that butyrate and to a lesser extent propionate transactivate Ppar γ [7]. To analyse butyrate-induced Ppar γ -dependent gene regulation, organoids were cultured from colonic crypt cells derived from Ppar $\gamma^{\Delta EEC}$ and WT mice, treated with butyrate and gene expression was measured using microarrays. To select Ppar γ -dependent genes regulated by butyrate, Venn plots were created from genes differentially regulated between butyrate vs. control in WT and Ppar $\gamma^{\Delta EEC}$ organoids, respectively (**Figure 6A**). From the differentially regulated genes in WT organoids, butyrate regulated 49% of genes Ppar γ dependently. A much higher number, however, was regulated by butyrate in the Ppar $\gamma^{\Delta EEC}$ derived organoids indicating that butyrate regulated much more genes related to Ppar γ deficiency.

To hypothesize on biological implications of butyrate-induced genes that are Ppar γ -dependent, gene sets significantly ($P < 0.001$, $FDR < 0.05$) enriched with genes regulated by butyrate in WT but not in Ppar $\gamma^{\Delta EEC}$ derived organoids were analysed. These gene sets encompassed metabolic, catabolic processes (lipid catabolism, fatty acid oxidation, TCA cycle, electron transport chain, redox reactions), Nrf2 targets, amino acid and glucose metabolism (**Figure 6B**) suggesting a role of butyrate and Ppar γ for energy-generating and oxidative stress response.

Ppar γ mediated inulin and butyrate-induced genes

We next compared the Ppar γ -dependent genes regulated with butyrate and IN. Genes only differentially regulated with butyrate (vs. control) in organoids derived from WT mice were selected and overlap with genes differentially regulated only in WT mice fed IN (vs. control) was calculated with Venn plots. The comparison revealed that only 6 out of 233 genes that were regulated by butyrate were also regulated in mice fed IN in Ppar γ dependent manner.

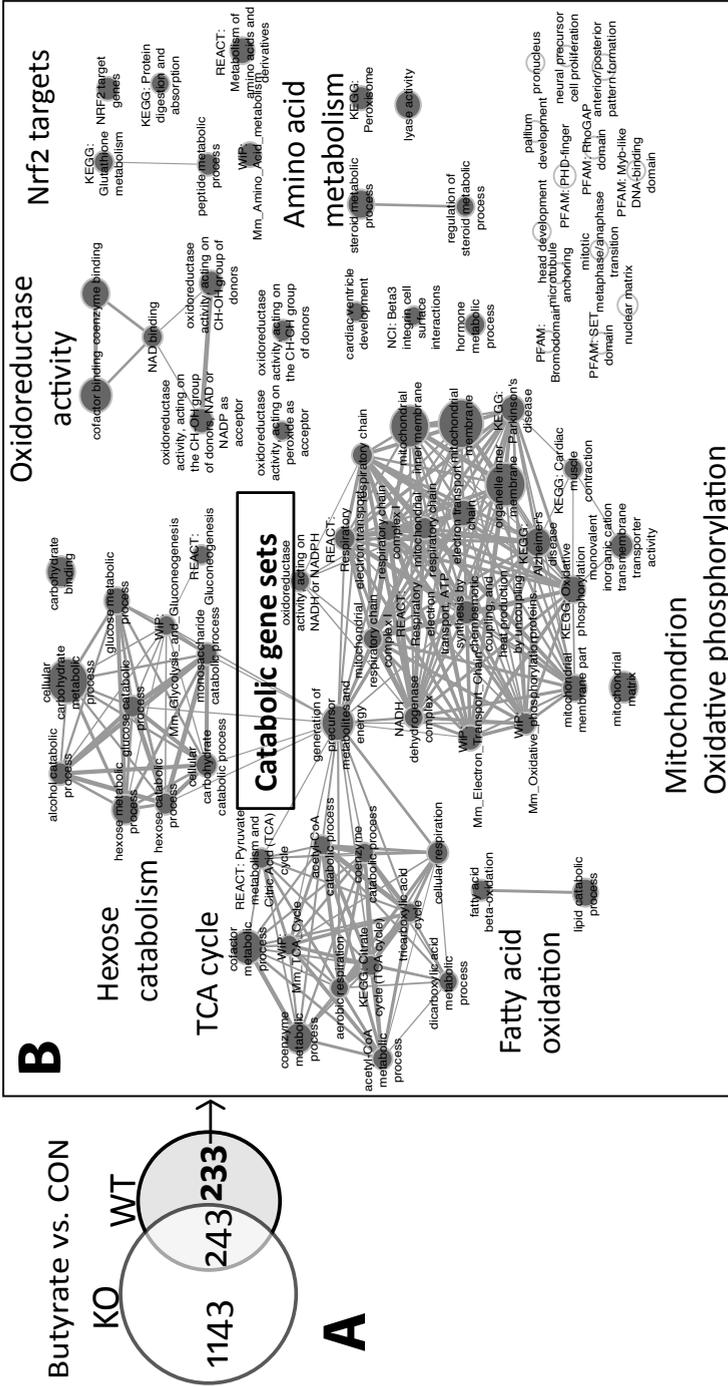


Figure 6

Comparison of significantly ($P < 0.01$) regulated genes between (A) butyrate and control (PBS) treated colonic organoids derived from WT (wild type) or KO (cre-vil *Pparγ* KO) mice; (B) gene sets significantly ($P < 0.001$, $FDR < 0.05$) enriched with genes only regulated in butyrate treated organoids derived from WT; dark grey indicates activated, white suppressed gene set



Discussion

Herein we report for the first time on the role of intestinal Ppar γ on gene expression regulation in colon of mice fed with the dietary fiber inulin. We demonstrated with whole-genome gene expression analysis that epithelial Ppar γ is associated with changes in genes related to metabolic, catabolic processes, cell cycle, and immune system in colon in response to fermentation of inulin. These results supports that epithelial PPAR γ plays an important role for signaling between epithelial cell and immune system in keeping mucosal homeostasis [16]. Furthermore, we used an organoid model derived from KO and WT mice treated with butyrate to compare effects induced by inulin. PPAR γ is transactivated by fiber fermentation products [7] but whole-genome gene expression profiles induced by butyrate in Ppar γ -dependent manner have not been reported before. Herein, we report that Ppar γ plays a role for gene regulation primarily related to energy metabolism and oxidative stress response (Nrf2) in butyrate treated colonic organoids. This result is in line with other studies showing that butyrate affect cellular energy metabolism by increasing fatty acid oxidation and mitochondrial respiration [18] and enhance phase II enzymes via Nrf2 [19]. Interplay and positive feedback loop between NRF2 and PPAR γ in regulating oxidative stress response and anti-inflammatory activities has been suggested before [20]. Although the observed overlap of butyrate- and IN-induced, Ppar γ -dependent genes was relatively small, biological processes associated with Ppar γ -dependent genes were largely comparable between IN and butyrate. These processes encompassed mainly energy generating processes and oxidative stress response (Nrf2). Although the expression changes in immune-related and Ppar γ -dependent genes suppressed in mice fed inulin cannot be compared with the organoid model, this does not exclude the possibility that SCFA can be responsible for part of the immune-related changes seen *in vivo* [17].

Regulation of SCFA transport via Ppar γ - Slc5a8

Notably, we observed higher integral levels for acetate, propionate and butyrate in KO mice fed inulin diet compared to WT mice fed control or inulin diet. This effect suggests that fermentation of dietary fiber and SCFA metabolism in epithelial, intestinal cells is blocked by abolishing Ppar γ because Ppar γ might have an effect on the absorption of SCFA, in particular butyrate. Butyrate is transported by Slc5a8/Smct1 into colonocytes (sodium-coupled monocarboxylate transporter 1) [21]. The mRNA levels of this SCFA transporter in cecum were significantly correlating with all three SCFA concentrations found in cecum in mice fed guar gum [22]. Gene expression of *Slc5a8* was indeed significantly higher in IN fed WT mice, however, not in KO mice (fold change 1.23, P=0.01 WT IN and 1.07, P=0.35 KO IN). Hence, the higher SCFA level in the KO mice fed inulin could indicate less uptake by Smct1 which might be dependent on intestinal, epithelial Ppar γ .

Taken together, we demonstrated that Ppar γ -dependent gene expression changes induced by butyrate are related to Nrf2 and energy metabolism in epithelial cells, which is comparable with the gene expression response seen in inulin fed mice. The abolished increase in SCFA transporter mRNA level in KO mice suggest a role of Ppar γ in uptake of SCFA which might support the importance of butyrate in regulating the transcriptional response to inulin via Ppar γ . Little overlap between IN and butyrate of the regulated Ppar γ -dependent genes might suggest that other metabolites or different cofactors additionally play a role in mediating response to fiber fermentation.

Acknowledgments

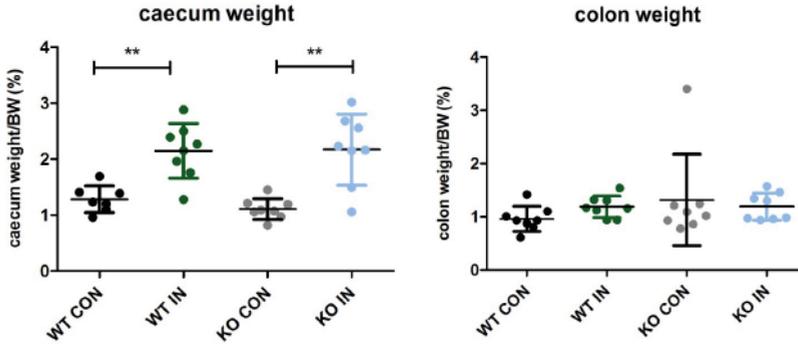
We would like to thank Dr Diederick Meyer (Sensus) for his kind gift of inulin, Kim Jannink, Shohreh Keshtkar and Jenny Jansen, Division of Human Nutrition, Wageningen University, for their skilful technical assistance.



References

1. Lim CC, Ferguson LR, Tannock GW (2005) Dietary fibres as "prebiotics": implications for colorectal cancer. *Mol Nutr Food Res* 49: 609-619.
2. Balakrishnan M, Floch MH (2012) Prebiotics, probiotics and digestive health. *Curr Opin Clin Nutr Metab Care* 15: 580-585.
3. Slavin JL (2013) Fiber and prebiotics: mechanisms and health benefits. *Nutrients* 5: 1417-1435.
4. Eswaran S, Muir J, Chey WD (2013) Fiber and functional gastrointestinal disorders. *Am J Gastroenterol* 108: 718-727.
5. Topping DL, Clifton PM (2001) Short-chain fatty acids and human colonic function: roles of resistant starch and nonstarch polysaccharides. *Physiol Rev* 81: 1031-1064.
6. Roediger WE (1982) Utilization of nutrients by isolated epithelial cells of the rat colon. *Gastroenterology* 83: 424-429.
7. Alex S, Lange K, Amolo T, Grinstead JS, Haakonsson AK, et al. (2013) Short-chain fatty acids stimulate angiopoietin-like 4 synthesis in human colon adenocarcinoma cells by activating peroxisome proliferator-activated receptor gamma. *Mol Cell Biol* 33: 1303-1316.
8. Tontonoz P, Spiegelman BM (2008) Fat and beyond: the diverse biology of PPARgamma. *Annu Rev Biochem* 77: 289-312.
9. McAlpine CA, Barak Y, Matise I, Cormier RT (2006) Intestinal-specific PPARgamma deficiency enhances tumorigenesis in ApcMin/+ mice. *Int J Cancer* 119: 2339-2346.
10. Mohapatra SK, Guri AJ, Climent M, Vives C, Carbo A, et al. (2010) Immunoregulatory actions of epithelial cell PPAR gamma at the colonic mucosa of mice with experimental inflammatory bowel disease. *PLoS One* 5: e10215.
11. Sato T, Vries RG, Snippert HJ, van de Wetering M, Barker N, et al. (2009) Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature* 459: 262-265.
12. He W, Barak Y, Hevener A, Olson P, Liao D, et al. (2003) Adipose-specific peroxisome proliferator-activated receptor gamma knockout causes insulin resistance in fat and liver but not in muscle. *Proc Natl Acad Sci U S A* 100: 15712-15717.
13. Haenen D, Souza da Silva C, Zhang J, Koopmans SJ, Bosch G, et al. (2013) Resistant starch induces catabolic but suppresses immune and cell division pathways and changes the microbiome in the proximal colon of male pigs. *J Nutr* 143: 1889-1898.
14. Sartor MA, Tomlinson CR, Wesselkamper SC, Sivaganesan S, Leikauf GD, et al. (2006) Intensity-based hierarchical Bayes method improves testing for differentially expressed genes in microarray experiments. *BMC Bioinformatics* 7: 538.
15. Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3: Article3.
16. Wittkopf N, Neurath MF, Becker C (2014) Immune-epithelial crosstalk at the intestinal surface. *J Gastroenterol* 49: 375-387.
17. Meijer K, de Vos P, Priebe MG (2010) Butyrate and other short-chain fatty acids as modulators of immunity: what relevance for health? *Curr Opin Clin Nutr Metab Care* 13: 715-721.
18. Donohoe Dallas R, Garge N, Zhang X, Sun W, O'Connell Thomas M, et al. (2011) The Microbiome and Butyrate Regulate Energy Metabolism and Autophagy in the Mammalian Colon. *Cell Metabolism* 13: 517-526.
19. Yaku K, Enami Y, Kurajyo C, Matsui-Yuasa I, Konishi Y, et al. (2012) The enhancement of phase 2 enzyme activities by sodium butyrate in normal intestinal epithelial cells is associated with Nrf2 and p53. *Mol Cell Biochem* 370: 7-14.
20. Polvani S, Tarocchi M, Galli A (2012) PPARgamma and Oxidative Stress: Con(beta) Catenating NRF2 and FOXO. *PPAR Res* 2012: 641087.
21. Miyauchi S, Gopal E, Fei YJ, Ganapathy V (2004) Functional identification of SLC5A8, a tumor suppressor down-regulated in colon cancer, as a Na(+)-coupled transporter for short-chain fatty acids. *J Biol Chem* 279: 13293-13296.
22. den Besten G, Havinga R, Bleeker A, Rao S, Gerding A, et al. (2014) The short-chain fatty acid uptake fluxes by mice on a guar gum supplemented diet associate with amelioration of major biomarkers of the metabolic syndrome. *PLoS One* 9: e107392.

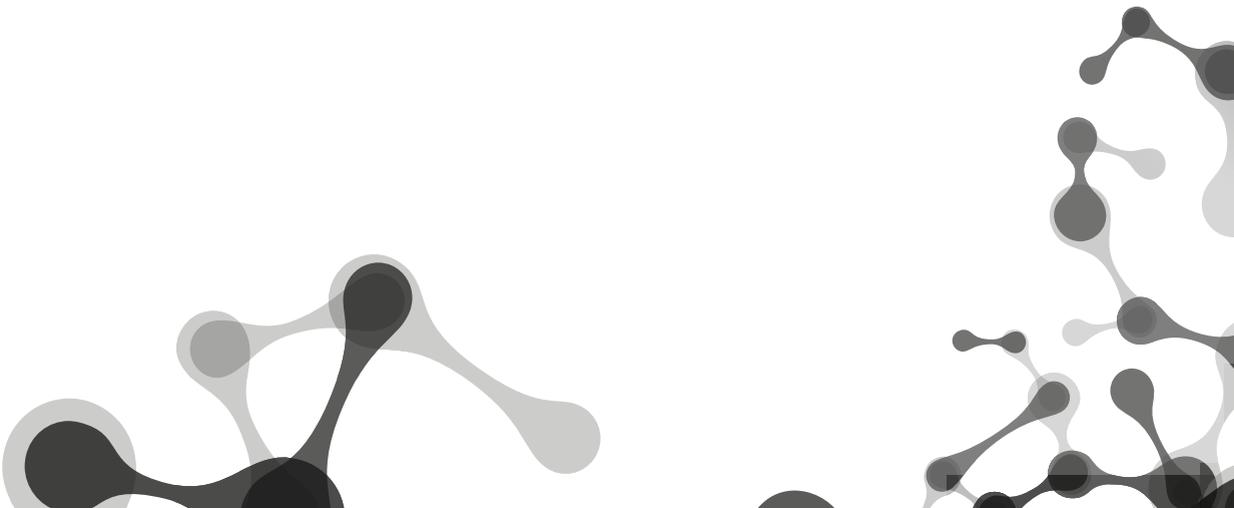
Supplemental Information



Supplemental Figure 1

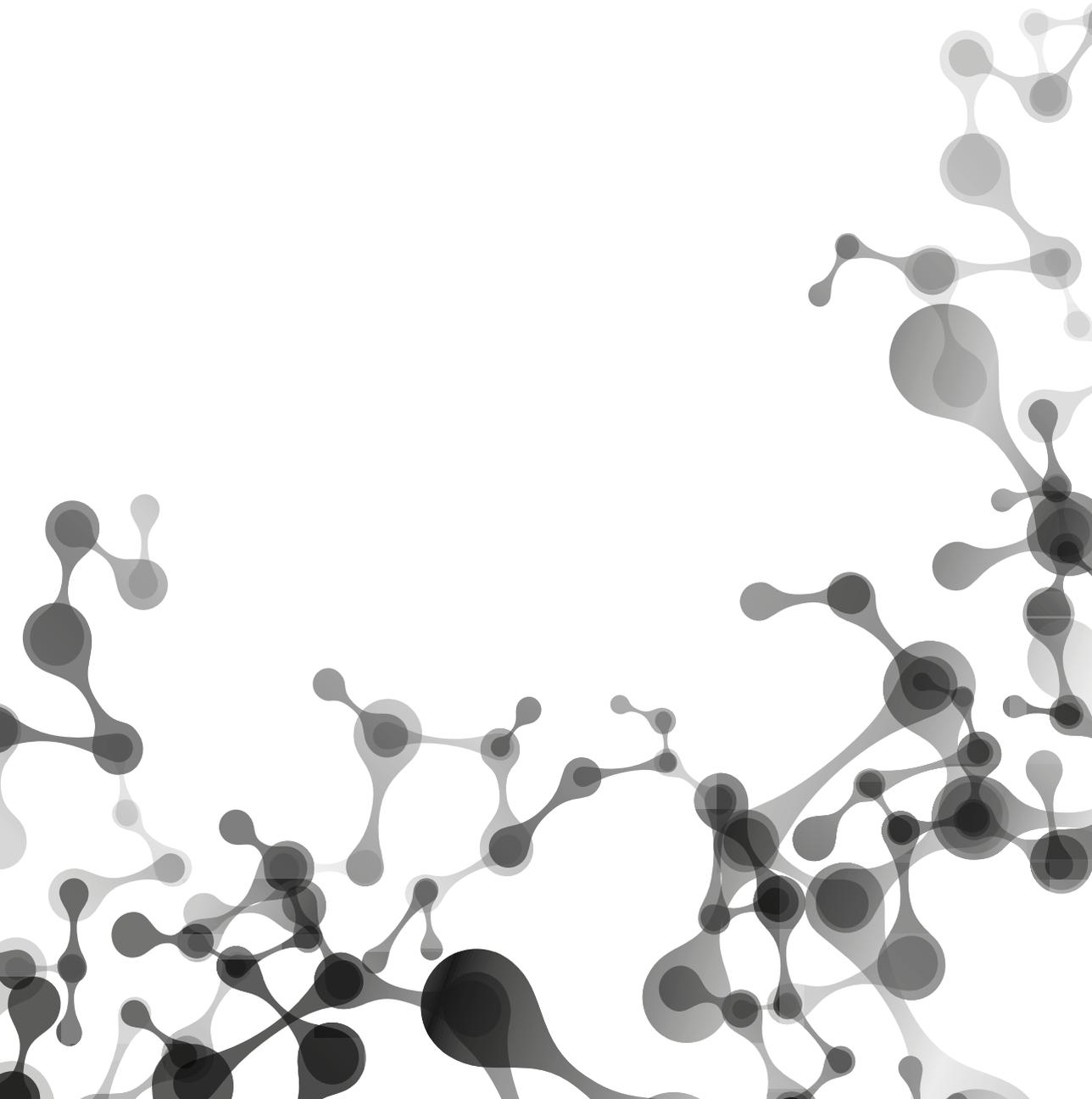
Weight of cecum and colon including content in mice after 10 days feeding inulin (IN) or control (CON) diet. Average \pm SD weight is given per group of mice fed inulin or control diet in wild type mice (WT) or cre-vil *Ppar γ* KO mice (KO)





Chapter 7

General Discussion



Health is not merely the ‘absence of disease’ (WHO, 1948), the definition should also include the ability to adapt and self-manage [1]. Intestinal health includes criteria such as nutrition, intestinal microbiota, immune status, absence of gastro-intestinal illness and well-being [2]. Hence, to prevent diseases, understanding the interactions between food/diet, microbiota, and host physiology is important [2].

Therefore, the work presented in this thesis aimed at characterizing the molecular relationships between dietary fibers and the large intestine, with focus on short-chain fatty acid-related transcriptional regulation in large intestine in mice. To comprehensively elucidate the underlying mechanisms, omics technologies were used to study both host and bacterial gene regulation and microbiota composition, which subsequently were integrated with multivariate statistics. To test the generated hypothesis, an intestine-specific knock out mouse models was used.

The following main results were obtained in this thesis:

1. SCFA –related gene expression changes in large intestine differ between a low fat/high carbohydrate and high fat/low carbohydrate diet (**chapter 2**)
2. Dietary fibers that are fermented induce a common transcriptional profile that is potentially governed by Ppar. The common gene expression changes were related to energy-generating processes and correlated with bacteria mainly belonging to *Clostridium* cluster XIVa, which encompass known butyrate producer. Furthermore, unique genes regulated by different dietary fibers were identified (**chapter 3**).
3. Next to commonalities induced by dietary fibers, metatranscriptome analysis of cecal microbiota revealed that different fibers activate different bacterial families, which are related to distinct activities in the fermentation of fibers into SCFA (**chapter 4**).
4. SCFA, in particular butyrate, (trans)activate PPAR γ and regulate ANGPTL4 in colonic cell line (**chapter 5**).
5. The transcriptional response to fermentation of inulin and butyrate is partially governed by epithelial Ppar γ in colon (**chapter 6**).

Ppar γ in the large intestine

Next to adipose tissue, Ppar γ is also highly expressed in the large intestine (<http://biogps.org>). While PPAR γ is known to play an important role in adipose tissue, e.g. for differentiation [3], the function of PPAR γ in colon, however, is less clear. Activation of Ppar γ by its ligand troglitazone has been seen to be protective on malignant changes involved in colon cancer [4]. Whereas others found increased formation of colon polyps upon Ppar γ activation by troglitazone and suggest that Ppar γ plays a role in mediating effects of high-fat diet-induced colorectal tumorigenesis [5]. Both SCFA and dietary fibers are related to decreased cancer risk, and are inducing Ppar γ targets (**chapter 3** and **5**). What the impact of high-fat diet is on protective effects of SCFA and dietary fibers via Ppar γ should

be investigated. The type of ligand might also play a role for effects of Ppar γ on colonic dysbiosis.

In addition, we noticed many genes that were regulated independent of epithelial Ppar γ in inulin fed mice in colon (**chapter 6**) indicating that other mechanisms may play a role including Ppar γ in other cell types or related to the dysfunction of epithelial Ppar γ as antagonist for other regulatory networks. In rosiglitazone treated cre-vil Ppar γ KO mice reduced colonic inflammation was seen and hence independent of epithelial Ppar γ [6]. Ppar γ is also expressed in other immune related cell types such as T and B cell, macrophages, eosinophils or dendritic cells [7]. Hence, it can be speculated that activation of Ppar γ expressed in other cell types might play a role for colonic response.

Epithelial vs. immune-cell Ppar γ in intestine

Studies with mice with experimental inflammatory bowel disease investigated the influence of immune cell and epithelial Ppar γ KO (knockout) on mucosal homeostasis. Studies with epithelial Ppar γ KO mice demonstrated a role of Ppar γ for mucosal immune response possibly via lysosomal and antigen presenting pathways [8]. Studies with T cell specific Ppar γ KO mice demonstrated regulation of genes in colonic mucosa that are related to adhesion dynamics, TCA cycle, ribosome and apoptosis [9]. Similarly, macrophage specific Ppar γ KO showed increased expression of fatty acid oxidation and inflammatory genes in colonic mucosa and an influence on T cells [10], demonstrating the impact and the complex interaction of Ppar γ in both epithelial and immune cell types for mucosal homeostasis.

Nrf2 and Ppar γ

Next to PPAR, Nrf2 (NF-E2-related factor 2) target genes were regulated in response to dietary fiber (**chapter 3**). Like Ppar γ , Nrf2 target genes were commonly increased for dietary fibers that are related to higher SCFA concentrations. Compared to Ppar γ KO mice, WT mice on control diet showed increased expression of Nrf2 target genes (GSEA, $P < 0.0001$, $FDR < 0.026$), indicating that Nrf2 acts in close relationship with PPAR γ . Nrf2 and Ppar γ are known to be in cross talk [11]. In addition, target genes of Nrf2 were not considered significantly regulated in inulin fed mice compared to control diet fed mice under KO conditions indicating that response to fermentation is also partially governed by Ppar γ -Nrf2. Taken together, response to dietary fiber fermentation in colon is likely not only governed by epithelial Ppar γ but is more complex and includes also other mechanisms that play an important role.

High-fat/low carb diet induced changes in large intestinal physiology

The transcriptional response to SCFA infusion was different between low fat/high carb (LFD) and high fat/low carb (HFD) diets (**chapter 2**). Therefore, high-fat diet induced changes in intestinal physiology will be discussed in the following paragraph.

Dietary fat is taken up in the small intestine. Hence, most studies on dietary fibers are rather focusing on the small intestine than large intestine. An excess of dietary fat as part of a high fat diet has influences on the intestinal physiology in particular in the more distal part [12] but also in the cecum and colon. The type of dietary fat also affects the fecal microbiota diversity and composition [13]. DeWit et al. [13] showed in this study that saturated fat (palm oil) deviated from all other dietary fat types in regulating gene expression in the distal part of the small intestine indicating an overflow of fat to the more distal parts after 2 weeks diet intervention. Hence, dietary fat might directly in even more distal parts, in large intestine, affect gut physiology. In line with this, it was shown that high fat diet induces changes in the large intestinal physiology (cecum), with respect to microbiota diversity, composition and metabolome as well as metaproteome [14]. Interestingly, it was demonstrated that changes in the diet induced more marked changes on the “chemical fingerprint” (unique metabolite pattern) than major perturbations in microbiota composition (e.g. by antibiotics treatment) which indicates the high impact of diet on intestinal physiological. The functional changes induced by HFD included shifts towards amino acid and steroid metabolism. The changes in amino acid metabolism with HFD might be due changes in ratio of protein to carbohydrates in HFD compared to LFD, which results in higher production of branched-chain fatty acids. In **chapter 2** we observed regulation of genes related to branched-chain amino acids and other amino acids after propionate infusion on HFD, which might indicate the interference of propionate with HFD metabolites and hence might contribute to positive effects of dietary fiber.

Furthermore, HFD are related to changes in bile acid metabolism induced by bacteria, which form secondary bile acids such as deoxycholic acid and lithocholic acid. In rat colon, deoxycholic acid and butyrate have antagonistic effects on proliferation along the crypt-villus axis [15]. Due to the lowering in pH by adding SCFA, the transformation of bile acids is inhibited [16] which might prevent cell damage and increased proliferative activity with detrimental health effects. Hence, possible bile acid related metabolites should be studied in relation to SCFA induced changes.

G-protein coupled receptor (Gpr) 41 are receptors for SCFA. SCFA, however, have been reported to protect against diet-induced obesity independent of Gpr41 after HFD feeding [17]. Hence, other mechanism might play a role. HFD induces microbiota changes leading to inflammation via LPS-induced TLR4 activation [18]. TLR signaling and PPAR signaling respond differentially to changes in microbial activity related to different dietary fiber interventions as was seen from correlations between metatranscriptome and host transcriptome (**chapter 4**). It has been shown that activation of macrophage-specific PPAR gamma signaling is inhibited by activation of TLR4 in LPS-stimulated macrophages [19]. It can be speculated that activation of PPAR γ by SCFA is related to and possibly counteracted by TLR4 signaling upon HFD feeding, a mechanism that should be further elaborated.

Taken together, the impact of the diet on the intestinal metabolome in influencing mucosal transcriptome response should be studied more comprehensively when investigating intestinal homeostasis.

Dietary fiber-induced changes mediated via SCFA?

Compared with the transcriptional profiles induced by SCFA, profiles induced by dietary fibers were different to SCFA infusions (mainly with respect to metabolic, catabolic gene sets) (**chapter 2, 3**). Several factors might play a role in this discrepancy. For example, it has been shown that individual SCFA can have opposite effects in metabolic pathways [20], [21]. Hence, the studied individual SCFA might act differently when applied in combination, as it is the case in more physiological conditions. Another point that can be taken into account for explaining differences between response to SCFA infusion and dietary fibers are fluxes of SCFA vs. snapshot concentrations. By measuring SCFA concentrations, production and uptake rates of SCFA remain unknown. It has been shown that actual concentrations of SCFA do less correlate with parameters of metabolic syndrome after feeding guar gum to mice than do flux measurements of SCFA [22]. Hence, different fibers can have similar SCFA concentrations but different uptake rates, which might lead to different effects.

Role of concentration dependent effect of butyrate

Butyrate in different concentrations induced different sets of genes (**chapter 5**; 1mM vs. 8mM) which points to other mechanism being involved at different concentrations. This is further supported by the fact that butyrate regulated also many genes *Ppary* independently (243 genes regulated in both *Ppary* WT and KO mice, **chapter 6**). It was shown that butyrate induced different mechanisms for concentrations > 5mM where HDAC inhibition takes place which is different at lower concentrations [23]. It is suggested that the different concentrations may play a role for cell turnover along the lumen-to-crypt axis of the colon [24]. At lower concentrations, which can be found at the crypt cells, butyrate induces proliferation of cells via usage as energy substrate and histone acetylation via acetyl-CoA. Whereas at the luminal site butyrate concentrations are higher, exceeding metabolic capacity of the cells and hence inducing apoptosis of cells via HDAC (histone deacetylase) inhibition. Hence, detailed mechanism about the different concentrations not only along the intestine, in small or large intestine, proximal or distal, but also along the crypt-villus axis should be studied in more detail. Experiments on the crypt or surface cells can reveal more evidence for the expected role of butyrate and also PPAR along the crypt-villus axis. Transcriptome analysis of crypt and surface cells isolated by laser capture microdissection can be used to dissect the molecular mechanism on the single cell level along crypt-villus axis [25].

Furthermore, other metabolites formed during dietary fiber fermentation and degradation might explain differences in colonic gene expression response.

Intestinal metabolites

In general, intestinal metabolites modulated or produced by microbiota include bile acid derivatives, indole, choline metabolites, phenolic, benzoyl, phenyl derivatives, vitamins, polyamines, lipids [26]. During fermentation of carbohydrates mainly SCFA are formed, but also formate, hydrogen sulphide, ethanol, succinate and lactate. In **chapter 6**, besides SCFA, another metabolite (succinate) was found to be important for driving the intestinal metabolome after feeding mice inulin. Succinate is a di-carboxylic acid and an intermediate product of fermentation which is converted into propionate by luminal bacteria [27]. Metabolites such as succinate are rather used in the metabolic network of the gut bacteria to finally produce SCFA, propionate [28]. Succinate can be taken up by large intestinal mucosa [29]. Little is known about effects on gene expression in intestinal mucosa by succinate. Succinate can be metabolized in the TCA cycle. Succinate was further shown to be a signaling molecule in inflammation [30], [31] and linking TCA cycle dysfunction to oncogenesis [32]. *In vitro* studies showed that succinate can inhibit growth and proliferation of colon cancer cells and that some polyphenols can increase the cecal succinate levels in rats fed HFD [33]. Other products derived from carbohydrate degradation are less well studied with respect to host metabolism.

Alternatives

What would be alternatives to study impact of SCFA in dietary fiber mediated effects? *Germ free* animal models can be applied to study the direct effects of dietary fibers on colonic mucosa without influences of SCFA, which are not produced in germ free mice. This model, however, has several marked disadvantages such as large differences between germ free and conventional mice with respect to immune cells, epithelial cell turnover, mucosal gene expression and metabolic functions [34] making it difficult to compare to physiological conditions. Also short-term *antibiotics* use to diminish the number of bacteria could be applied to study effects of dietary fiber without SCFA production [35]. A shift in the balance of microbial ecosystem [36], however, might influence the response to dietary fibers in the colonic epithelium. Furthermore, to test the causal effect of SCFA in dietary fiber induced gene expression profile one could use SCFA transporter knockout models. SCFA are taken up via transport proteins (monocarboxylic acid transporters), mainly by MCT1-4. Also sodium-coupled transporter SMCT1 has been found to transport SCFA. Fluxes of SCFA correlate with mRNA level of these transporters [22] implicating the importance of these transporters in uptake and subsequent metabolism of SCFA. Hence, blocking uptake of SCFA via knock out of these transporters might help to better understand the impact of intracellular SCFA metabolism. Next to SCFA, the monocarboxylate transporter also transport other metabolites (lactate, pyruvate, nicotinate and ketone

bodies) [37]. Knock down of MCT1 and MCT4 by siRNA has been seen to lower both lactate and butyrate uptake into rat intestinal cells [35]. From these metabolites, lactate and pyruvate are formed as intermediary products by bacteria during carbohydrate fermentation, which are utilized by other bacteria to form SCFA. Hence, uptake of these metabolites by these transporters might play only a minor role for host response to fermentation. Besides uptake via transporters, however, diffusion of the protonated form of SCFA has also been suggested [38] and the primary transporter for butyrate on the apical site has not been unequivocally identified yet. Nevertheless, for answering the question in as how much SCFA are mediating the colonic response to fermentation of dietary fiber, *Smct1* knock out might be most appealing since in **chapter 6** we showed that mRNA level were significantly increased by dietary fibers in Ppar γ dependent manner and might hence be important for mediating effects of dietary fiber.

Special case of resistant starch

Resistant starches are starches included in the definition of dietary fiber since they are not digested in the small intestine and can be fermented in the large intestine. There are four types of resistant starch, type 1 (physical inaccessible), 2 (ungelatinized starch), 3 (retrograded starch) and 4 (chemically modified starch). The RS used in our studies was tapioca starch, type 3. The same dietary fiber has been used previously in pigs fed RS for 14 days [39]. In this study, the colonic gene expression changes to RS in pigs showed a comparable profile to mice fed inulin, FOS, guar gum or arabinoxylan supplemented diets fed for 10 days, but not to those of mice fed RS (**chapter 3**). In both cases, genes related to energy metabolism were differentially regulated and PPAR γ was identified as potential upstream regulator. Similarly, SCFA levels were increased with RS in pigs as with the other fiber diets in mice. Both the gene expression changes and luminal SCFA level in RS fed mice were rather similar to control conditions in the mouse study. It can be speculated that in the mouse model RS is less efficiently fermented to SCFA and hence comparable to normal starch that still is present in the control diet condition. It was suggested that some obese mouse models do not ferment RS [40]. In humans, in a randomised, placebo controlled trial RS (type 2) was tested in carrier of hereditary colorectal cancer (Lynch syndrome) for 4 years. As a result, no protection of development of colorectal cancer was found with RS [41]. Similar no effect results were found by [42] and [43]. In the latter study, fecal total SCFA concentrations were measured and showed no change compared to control with RS2 or RS3. In conclusion, the results show that dietary fibers can act quite differently in different animal models and human studies.

Which dietary fiber is the best?

The answer to this question depends on what is healthy. Production of SCFA is generally considered beneficial for gut health by serving energy to colonic mucosa, hence IN and GG-enriched diets are probably most healthy. The actual SCFA concentrations, however,

are less representative for host response as different uptake rates can be expected. From the definition mentioned before, studying gut health necessarily includes nutrition, intestinal microbiota and immune status. By comprehensively studying diet, microbiota and host response we identified potential mechanism related to health. FOS and AX were quite comparable in gene expression profile to IN and GG in activating PPAR target genes which is considered healthy for gut and hence health-related effects might be attributed to these fibers, too. Next to the common overlapping effects of fermentable fibers, we saw unique effects in gene regulation. What the consequences of the fine and unique effects of specific dietary fiber are in preventing disease should be studied in long-term studies and in combination with knock out models of specific, potential upstream regulators identified in **chapter 3**. Health definition should also include ability to adapt, self-manage [1]. From that perspective, however, we can speculate that a combination of different fibers might be even better since many unique genes are regulated by individual fibers possibly enhancing adaptability.

Conclusion and Future direction

The definition of dietary fiber is the bottleneck in understanding its health benefits. As mentioned before, the definition includes complex carbohydrates and lignin that are not fermented in small intestine and can be fermented in large intestine. Considering that a variety of sources of dietary fibers with different chemical compositions exist, a wide array of specific and unique fiber-related microbial metabolites can be expected related to the identified specific genes and should be studied with respect to host metabolism. We showed that the diversity of dietary fibers can be studied comprehensively with omics tools to better understand these important food components. The plethora of other potential metabolites conferring health benefits, however, needs to be explored further. Metabolomics should be included when studying diet-host-microbe interactions as diet clearly has a large impact on the intestinal metabolome, which in turn is important for understanding host dynamics. Finally, to better understand the biological basis of the interactions of multi-omics data, one needs to test the generated hypotheses in subsequent studies.

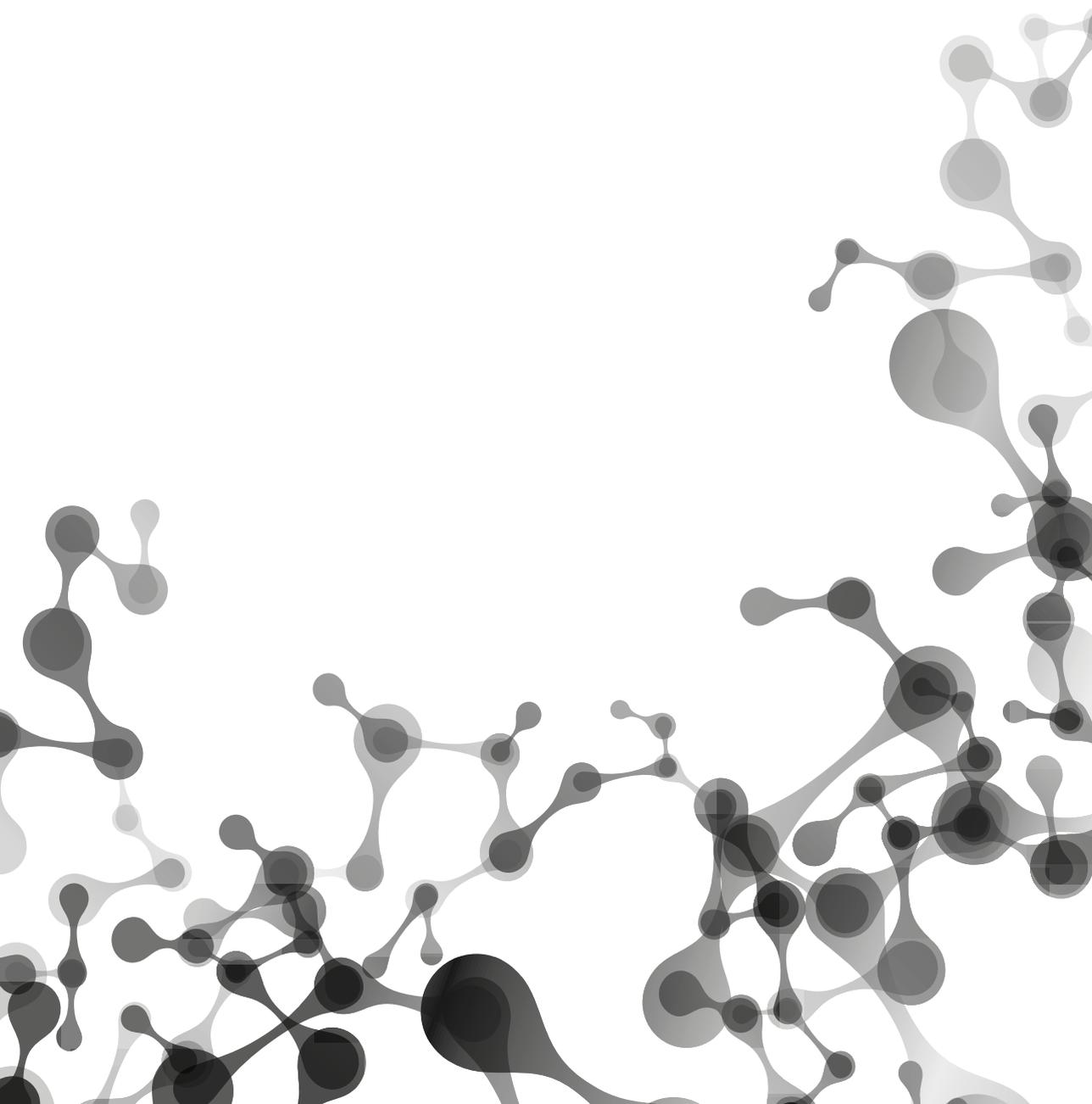
References

1. Huber M, Knottnerus JA, Green L, van der Horst H, Jadad AR, et al. (2011) How should we define health? *BMJ* 343: d4163.
2. Bischoff SC (2011) 'Gut health': a new objective in medicine? *BMC Med* 9: 24.
3. Ahmadian M, Suh JM, Hah N, Liddle C, Atkins AR, et al. (2013) PPARgamma signaling and metabolism: the good, the bad and the future. *Nat Med* 19: 557-566.
4. Sarraf P, Mueller E, Jones D, King FJ, DeAngelo DJ, et al. (1998) Differentiation and reversal of malignant changes in colon cancer through PPARgamma. *Nat Med* 4: 1046-1052.
5. Saez E, Tontonoz P, Nelson MC, Alvarez JG, Ming UT, et al. (1998) Activators of the nuclear receptor PPARgamma enhance colon polyp formation. *Nat Med* 4: 1058-1061.
6. Adachi M, Kurotani R, Morimura K, Shah Y, Sanford M, et al. (2006) Peroxisome proliferator activated receptor gamma in colonic epithelial cells protects against experimental inflammatory bowel disease. *Gut* 55: 1104-1113.
7. Szeles L, Torocsik D, Nagy L (2007) PPARgamma in immunity and inflammation: cell types and diseases. *Biochim Biophys Acta* 1771: 1014-1030.
8. Mohapatra SK, Guri AJ, Climent M, Vives C, Carbo A, et al. (2010) Immunoregulatory actions of epithelial cell PPAR gamma at the colonic mucosa of mice with experimental inflammatory bowel disease. *PLoS One* 5: e10215.
9. Guri AJ, Mohapatra SK, Horne WT, 2nd, Hontecillas R, Bassaganya-Riera J (2010) The role of T cell PPAR gamma in mice with experimental inflammatory bowel disease. *BMC Gastroenterol* 10: 60.
10. Hontecillas R, Horne WT, Climent M, Guri AJ, Evans C, et al. (2011) Immunoregulatory mechanisms of macrophage PPAR-gamma in mice with experimental inflammatory bowel disease. *Mucosal Immunol* 4: 304-313.
11. Polvani S, Tarocchi M, Galli A (2012) PPARgamma and Oxidative Stress: Con(beta) Catenating NRF2 and FOXO. *PPAR Res* 2012: 641087.
12. de Wit NJ, Boekschoten MV, Bachmair EM, Hooiveld GJ, de Groot PJ, et al. (2011) Dose-dependent effects of dietary fat on development of obesity in relation to intestinal differential gene expression in C57BL/6J mice. *PLoS One* 6: e19145.
13. de Wit N, Derrien M, Bosch-Vermeulen H, Oosterink E, Keshtkar S, et al. (2012) Saturated fat stimulates obesity and hepatic steatosis and affects gut microbiota composition by an enhanced overflow of dietary fat to the distal intestine. *Am J Physiol Gastrointest Liver Physiol* 303: G589-599.
14. Daniel H, Moghaddas Gholami A, Berry D, Desmarchelier C, Hahne H, et al. (2014) High-fat diet alters gut microbiota physiology in mice. *ISME J* 8: 295-308.
15. Velazquez OC, Seto RW, Bain AM, Fisher J, Rombeau JL (1997) Deoxycholate inhibits *in vivo* butyrate-mediated BrDU labeling of the colonic crypt. *J Surg Res* 69: 344-348.
16. Nagengast FM, Hectors MP, Buys WA, van Tongeren JH (1988) Inhibition of secondary bile acid formation in the large intestine by lactulose in healthy subjects of two different age groups. *Eur J Clin Invest* 18: 56-61.
17. Lin HV, Frassetto A, Kowalik EJ, Jr, Nawrocki AR, Lu MM, et al. (2012) Butyrate and propionate protect against diet-induced obesity and regulate gut hormones via free fatty acid receptor 3-independent mechanisms. *PLoS One* 7: e35240.
18. Kim KA, Gu W, Lee IA, Joh EH, Kim DH (2012) High fat diet-induced gut microbiota exacerbates inflammation and obesity in mice via the TLR4 signaling pathway. *PLoS One* 7: e47713.
19. Necela BM, Su W, Thompson EA (2008) Toll-like receptor 4 mediates cross-talk between peroxisome proliferator-activated receptor gamma and nuclear factor-kappaB in macrophages. *Immunology* 125: 344-358.
20. Wolever TM, Spadafora P, Eshuis H (1991) Interaction between colonic acetate and propionate in humans. *Am J Clin Nutr* 53: 681-687.
21. Demigne C, Yacoub C, Morand C, Remesy C (1991) Interactions between propionate and amino acid metabolism in isolated sheep hepatocytes. *Br J Nutr* 65: 301-317.
22. den Besten G, Havinga R, Bleeker A, Rao S, Gerding A, et al. (2014) The short-chain fatty acid uptake fluxes by mice on a guar gum supplemented diet associate with amelioration of major biomarkers of the metabolic syndrome. *PLoS One* 9: e107392.
23. Donohoe DR, Collins LB, Wali A, Bigler R, Sun W, et al. (2012) The Warburg effect dictates the mechanism of butyrate-mediated histone acetylation and cell proliferation. *Mol Cell* 48: 612-626.

24. Bultman SJ (2014) Molecular pathways: gene-environment interactions regulating dietary fiber induction of proliferation and apoptosis via butyrate for cancer prevention. *Clin Cancer Res* 20: 799-803.
25. IJssennagger N, Derrien M, van Doorn GM, Rijnierse A, van den Bogert B, et al. (2012) Dietary heme alters microbiota and mucosa of mouse colon without functional changes in host-microbe cross-talk. *PLoS One* 7: e49868.
26. Nicholson JK, Holmes E, Kinross J, Burcelin R, Gibson G, et al. (2012) Host-gut microbiota metabolic interactions. *Science* 336: 1262-1267.
27. Wolin MJ (1981) Fermentation in the rumen and human large intestine. *Science* 213: 1463-1468.
28. Fischbach MA, Sonnenburg JL (2011) Eating for two: how metabolism establishes interspecies interactions in the gut. *Cell Host Microbe* 10: 336-347.
29. Wolffram S, Badertscher M, Scharrer E (1994) Carrier-Mediated Transport Is Involved in Mucosal Succinate Uptake by Rat Large-Intestine. *Experimental Physiology* 79: 215-226.
30. Mills E, O'Neill LA (2014) Succinate: a metabolic signal in inflammation. *Trends Cell Biol* 24: 313-320.
31. Tannahill GM, Curtis AM, Adamik J, Palsson-McDermott EM, McGettrick AF, et al. (2013) Succinate is an inflammatory signal that induces IL-1beta through HIF-1alpha. *Nature* 496: 238-242.
32. Selak MA, Armour SM, MacKenzie ED, Boulahbel H, Watson DG, et al. (2005) Succinate links TCA cycle dysfunction to oncogenesis by inhibiting HIF-alpha prolyl hydroxylase. *Cancer Cell* 7: 77-85.
33. Haraguchi T, Kayashima T, Okazaki Y, Inoue J, Mineo S, et al. (2014) Cecal succinate elevated by some dietary polyphenols may inhibit colon cancer cell proliferation and angiogenesis. *J Agric Food Chem* 62: 5589-5594.
34. Fritz JV, Desai MS, Shah P, Schneider JG, Wilmes P (2013) From meta-omics to causality: experimental models for human microbiome research. *Microbiome* 1: 14.
35. Yap IK, Li JV, Saric J, Martin FP, Davies H, et al. (2008) Metabonomic and microbiological analysis of the dynamic effect of vancomycin-induced gut microbiota modification in the mouse. *J Proteome Res* 7: 3718-3728.
36. Willing BP, Russell SL, Finlay BB (2011) Shifting the balance: antibiotic effects on host-microbiota mutualism. *Nat Rev Microbiol* 9: 233-243.
37. Ganapathy V, Thangaraju M, Gopal E, Martin PM, Itagaki S, et al. (2008) Sodium-coupled monocarboxylate transporters in normal tissues and in cancer. *AAPS J* 10: 193-199.
38. Charney AN, Micic L, Egnor RW (1998) Nonionic diffusion of short-chain fatty acids across rat colon. *Am J Physiol* 274: G518-524.
39. Haenen D, Souza da Silva C, Zhang J, Koopmans SJ, Bosch G, et al. (2013) Resistant starch induces catabolic but suppresses immune and cell division pathways and changes the microbiome in the proximal colon of male pigs. *J Nutr* 143: 1889-1898.
40. Zhou J, Martin RJ, Tulley RT, Raggio AM, Shen L, et al. (2009) Failure to ferment dietary resistant starch in specific mouse models of obesity results in no body fat loss. *J Agric Food Chem* 57: 8844-8851.
41. Mathers JC, Movahedi M, Macrae F, Mecklin JP, Moeslein G, et al. (2012) Long-term effect of resistant starch on cancer risk in carriers of hereditary colorectal cancer: an analysis from the CAPP2 randomised controlled trial. *Lancet Oncol* 13: 1242-1249.
42. Burn J, Bishop DT, Mecklin JP, Macrae F, Moeslein G, et al. (2008) Effect of aspirin or resistant starch on colorectal neoplasia in the Lynch syndrome. *N Engl J Med* 359: 2567-2578.
43. Heijnen ML, van Amelsvoort JM, Deurenberg P, Beynen AC (1998) Limited effect of consumption of uncooked (RS2) or retrograded (RS3) resistant starch on putative risk factors for colon cancer in healthy men. *Am J Clin Nutr* 67: 322-331.



Nederlandse samenvatting



Nederlandse samenvatting

De gezondheid van de darm wordt bepaald door interactie tussen de voeding, de darmbacterien (microbiota) en de gastheer. Voedingsvezels staan erom bekend dat ze de darmgezondheid positief beïnvloeden. Voedingsvezels zijn de eetbare delen van planten en bestaan uit complexe koolhydraten en lignine. Een karakteristiek kenmerk van voedingsvezels is dat zij niet verteerd kunnen worden in de dunne darm. Er bestaan een groot aantal voedingsvezels die verschillen in koolhydraatsamenstelling en wijze van afbraak in darm. De afbraak van de voedingsvezels in de darm vindt plaats in het ceacum en colon door darmbacteriën. De afbraak van de voedingsvezels in de darm resulteert voornamelijk in de productie van metabolieten zoals korte-keten vetzuren (KKVZ) acetaat, propionaat en butyraat. KKVZ kunnen vervolgens door de gastheer worden opgenomen. Butyraat wordt voornamelijk gebruikt door de epitheelcellen en is belangrijk voor energiemetabolisme van de darmcellen. Daarnaast vermindert butyraat het risico op dikke darmkanker. Dit effect op dikke darmkanker wordt daarom gezien als één van de positieve effecten van voedingsvezels op de gezondheid. Propionaat en acetaat worden opgenomen en via het bloed getransporteerd naar andere organen van de gastheer. Propionaat wordt in de lever gebruikt als substraat voor glucose in lever en acetaat wordt verder getransporteerd naar anderen organen. Darmbacteriën spelen een belangrijke rol voor de gastheer. Onderzoekers schatten dat de mens uit 10 keer meer microbiële cellen bestaat dan uit lichaamscellen, wat de rol van de bacteriën in het lichaam verder onderstreept. Darmbacteriën beïnvloeden het energie metabolisme van de gastheer. Het is bijvoorbeeld aangetoond dat het eiwit Angptl4 die effecten van darmbacteriën op energiemetabolisme beïnvloed. In hoeverre KKVZ de expressie van Angptl4 beïnvloeden is nog niet duidelijk. Deze, net genoemde achtergrondinformatie staat gedetailleerd beschreven in **hoofdstuk 1**.

Het doel van het onderzoek beschreven in dit proefschrift was om de mechanismen te ontrafelen die achter de effecten van voedingsvezels en KKVZ op de gastheer zitten. Daarbij hebben wij gebruik gemaakt van verschillende zogenoemde -omics technologieën. Deze -omics technologieën omvatten transcriptomics (het meten van de activiteit van alle genen genactiviteit middels mRNA van de gastheer), metatranscriptomics (het meten van de activiteit van alle genen in de darmbacteriën) en het meten van samenstelling van de bacteriën in de darm.

In **hoofdstuk 2** wordt beschreven wat de invloed is van KKVZ op de genexpressie in het colon. De KKVZ zijn in dit hoofdstuk rechtstreeks ingebracht in de dikke darm en de respons van de gastheer is bestudeerd middels de microarray technologie, wat een vorm van transcriptomics is. Door middel van deze experimenten hebben wij genen geïdentificeerd die enkel in expressie veranderen bij toediening van of acetaat, of propionaat of butyraat. In tegenstelling tot acetaat en butyraat die met name de activiteit van genen gerelateerd

aan vetmetabolisme veranderen, beïnvloedt propionaat juist genen die gerelateerd zijn aan het koolhydraat- en aminozuurmetabolisme. Verder hebben KKVZ een effect op diverse processen gerelateerd aan het metabolisme en de celcyclus. Tevens hebben wij onderzocht wat de invloed is van een dieet dat een laag of een hoog vetgehalte bevatte op de effecten van KKVZ. Hieruit bleek dat de invloed van propionaat op de genexpressie in de dikke darm grotendeels wordt bepaald door het dieet, maar voor acetaat en butyraat was dat niet het geval.

Onderzoek beschreven in **hoofdstuk 3** bestaat uit gedetailleerde analyses van de moleculaire effecten van verschillende voedingsvezels op de dikke darm. Uit het onderzoek bleek dat de voedingsvezels die worden afgebroken in de dikke darm en een verhoogde KKVZ concentratie in de dikke darm veroorzaken, vergelijkbare effecten hebben op de genexpressie profielen in de darmwand en de samenstelling van de darmbacteriën. De genen die veranderden door deze voedingsvezels waren gerelateerd aan energie metabolisme. Tevens correleerden deze genen met bacteriën die voornamelijk horen bij de *Clostridium* klasse XIVa. Uit de analyses bleek verder dat de voedingsvezels die worden afgebroken in de dikke darm een gemeenschappelijke genexpressieregulator hadden, namelijk de peroxisome proliferator-activated receptor (PPAR) γ . Tenslotte hebben wij ook door middel van verschillende -omics technologieën aangetoond dat de verschillende voedingsvezels unieke genexpressie profielen induceren, wat de diversiteit van de verschillende voedingsvezels onderstreept.

In **hoofdstuk 4** hebben wij onze resultaten van hoofdstuk 3 verder onderzocht. In dit hoofdstuk hebben wij naast de samenstelling van de darmbacteriën ook de activiteit van de darmbacteriën gemeten. Dit hebben wij gedaan door het sequensen van monsters uit de caecum-inhoud van muizen die behandeld waren met de verschillende voedingsvezels uit hoofdstuk 3. Uit deze analyse bleek dat verschillende soorten bacteriën deelnemen aan de afbraak van voedingsvezels. Deze voedingsvezels zorgden allemaal voor een specifieke verhouding KKVZ.

Hoofdstuk 5 beschrijft het onderzoek naar de regulatie van het eiwit Angptl4. Angptl4 is een mogelijke mediator tussen de darmbacteriën en het energie metabolisme van de gastheer. Uit het onderzoek in hoofdstuk 5 bleek dat Angptl4, dat wordt gereguleerd door PPAR, en wordt geactiveerd door blootstelling aan KKVZ in darmcellen. Tevens bleek uit deze studie dat de KKVZ vooral de PPAR isoform PPAR γ activeert.

Vervolgens wordt in **hoofdstuk 6** beschreven welke invloed Ppar γ in de darmcellen heeft op de afbraak van de voedingsvezel inuline. Om dit te bestuderen zijn knock-out muizen ontwikkeld die de transcriptie factor Ppar γ missen in de epitheelcellen in de darm. Net als in hoofdstuk 3/4 hebben deze muizen 10 dagen inuline gegeten. Vervolgens is de

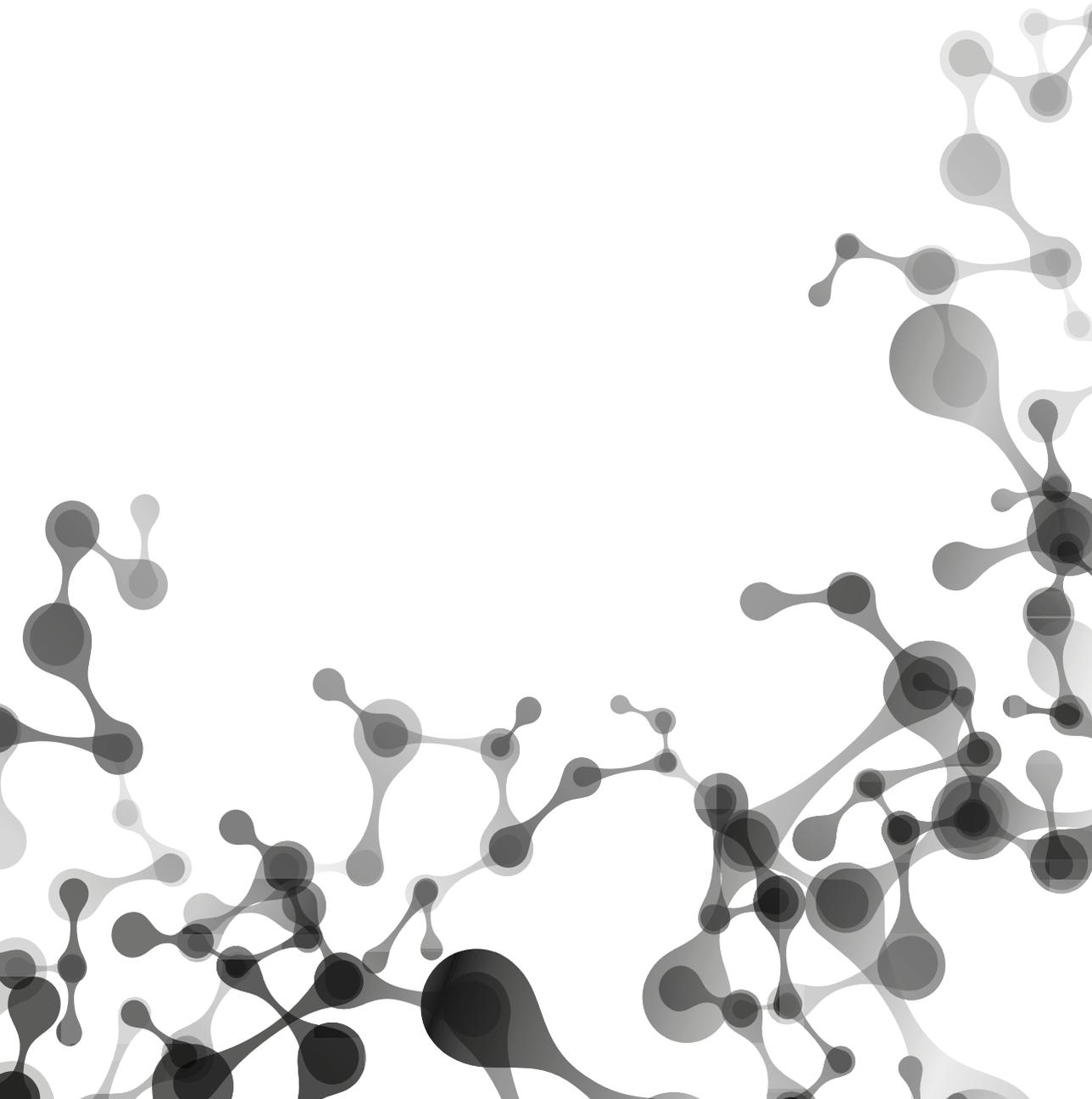
samenstelling van de darmbacteriën bepaald en de KKVZ concentraties en gen expressie in het darmepitheel gemeten. Uit deze analyses bleek dat de samenstelling van de darmbacteriën en het darm genexpressie profiel meer verschilde tussen de diëten dan tussen de genotypes. Verder hebben wij bepaald welke soorten darmbacteriën en welke genen beïnvloedt worden door Ppar γ . Hieruit bleek dat de Ppar γ afhankelijke genexpressieveranderingen gerelateerd aan de afbraak van inuline bestaan uit genen gerelateerd aan processen zoals immuun response, celcyclus, metabolisme, en oxidatieve stress.. Vervolgens hebben wij cellen geïsoleerd uit knock-out en wildtype muizen om zogenoemde darm-organoids ('mini-darm' die gekweekt is in het lab) te maken. Deze organoids werden daarna met butyraat behandeld en de genexpressie werd bepaald. Hieruit bleek dat butyraat een deel van Ppar γ afhankelijke genexpressie beïnvloedt. Deze genen waren gerelateerd aan het energie metabolisme, cel cyclus, DNA herstel, en oxidatieve stress.

In **hoofdstuk 7** worden de resultaten uit hoofdstukken 2 tot en met 6 bediscussieerd.

Op basis van de resultaten van dit proefschrift wordt geconcludeerd dat dat voedingsvezels en KKVZ zowel gemeenschappelijke en uniek effecten hebben op de profielen die wij hebben verkregen door middel van -omics technieken. Deze genexpressie en darmbacterie-profielen zijn cruciaal voor de regulatie van de darmfysiologie. Tevens is gebleken dat de effecten van KKVZ op het darmepitheel sterk afhankelijk zijn van het dieet waaruit de KKVZ worden geproduceerd en dat Ppar γ waarschijnlijk een regulerende rol heeft in de regulatie van de gen expressie responses van de gastheer.



Acknowledgement



The acknowledgement, het dankwoord, die Danksagung

It would be a pity to end with German negativity: "Da steh ich nun ich armer Tor und bin so klug als wie zu vor" (Goethe), I actually learned a lot the last 4 years. Although the Dutch seem more optimistic with "Alles komt goed", that belief alone would not make a thesis. Hence, I would like to thank a number of people whose input and influence made the thesis possible:

First of all, I would like to thank my promotor and co-promotor. Michael, danke, dass du mich als PhD Student angenommen und die Möglichkeit zur Promotion gegeben hast. Danke für alle konstruktiven Diskussionen und viele wunderbare Ideen. Ich habe deine kritischen Fragen bei unseren meetings (wenn auch manchmal im nachhinein) immer als Motivation gesehen, um noch mehr heraus zu holen und tiefer zu graben. Es gibt noch viel zu lernen und tun! Guido, as my daily supervisor, I would like to thank you for all your efforts the last 4,5 years. Thanks for all your enthusiasm for new experiments, data analyses strategies, for your help at all times (literally) and especially perseverance with manuscripts/ revisions/re...revisions. It was a colorful experience to work with you. You are full of ideas and possible solutions. There's always a way!

I would like to thank the committee, Jerry Wells, Koen Venema, Bert Groen and John Mathers for the critical review and judgment of the thesis!

Of course, I would like to thank the NCSB/TIFN project team. Hauke and Michiel, thanks for the discussions, help and many feedbacks with the manuscripts. Aat, thanks for organizing all the TIFN project meetings! Gijs and Dirk-Jan, our projectpartner from Groningen, thanks for the nice collaboration on the SCFA's and all the discussions during the project meetings.

Sander, thanks for the opportunity to work with you. I learned some nice techniques in the lab and I am happy and very grateful about the beautiful chapter! Frits and Sheril, I am thankful that you were teaching and helping me with the cell experiments, I learned a lot from you! Aafke thanks for your help with SCFA measurements. I am sure you will have great results; I wish you all the best!

I appreciate the help of many other people who made the experiments possible. Hans and Marcela from Groningen, thanks for your efforts with the PPARG KO mice. Great that we could collaborate! Henk and Melliana, from food chemistry, thanks for the saccharide analyses! Jacques, bedankt voor de interessante gesprekken over van alles en je hulp bij de NMR metingen. Noortje, bedankt voor al je werk met de organoids. Het was prettig om met jou samen te werken!

Mark and Philip, I am thankful for your help with R scripts, and all the questions around the microarray analysis!

I would also like to thank the students that helped with experiments, Felicia, Angie and Kim! It was fun working with you and good luck with your future plans!

Bert, Judith, Wilma en Rene, thanks for the help with the mouse experiments!

Thanks to Carolien, Karin, Mieke and Mechteld for all your help in the lab. Jenny and Shohreh, thanks for all the preparations for the many microarrays!

Shohreh, I am grateful for all your technical support and teaching, not only in Wageningen, but even in Groningen! You did a great amount of the lab work. You have a lot of drive and it was certainly fun working with you!

Carolien, bedankt voor de gezelligheid in Myanmar en de etentjes. Ook was het klimmen en hardlopen met jou heel prettig. Ik hoop dat we nog een keer samen de "berg" op kunnen lopen!

Danielle, naast de gezellige tijd in Myanmar, was het ook erg leuk om met jou samen te werken. Door jouw plezier aan structuur en planning was het een heel aangenaam en rustgevende samenwerking! Ik wens je veel succes met je toekomst plannen.

I would like to thank all the other NMgers/Pharmagroup for the great atmosphere at work, the uitjes, the NWO dagen and NUGO weeks. It was an unforgettable experience. And certainly, the NUGO week Munich was more than 4 beers ;-)! Mark, Michiel, Natasha, Mechteld, Diederik, Noortje and Pierluigi thanks for the gezelligheid and the friday borrels, it was a lot of fun with you in Wageningen. Lydia, Wilma, Rinke bedankt voor jullie interesse in het project. Parastoo, Wieneke, Lily, Fenni, Ya, Song, Juri en Sophie, good luck with your PhD!

Danke dem deutschen Stammtisch, Jessica, Mona, Susan&Susan, für bierstarke Unterhaltung und grenzüberschreitende Fahrten! Jessica, danke insbesondere für das Auffangen am Beginn. Dank deiner sozialen, fröhlichen und aktiven Art hatten wir viel Spaß und haben viel erlebt in 3 Jahren Wageningen. Danke auch für deine Hilfe, Interesse und kritische Diskussionen zu meinem PhD Projekt. Du warst ein wichtiger Teil, um es zu einer unvergesslichen Zeit zu machen!

Und Saskia, beinahe deutsch ;-), danke auch an dich fürs Zuhören und die schöne Zeit!

It would have been boring without my roommates, Neeraj and The Pink Ladies, Inge, Jvalini and Nikkie. Thanks for all the great time. I enjoyed the many chocolates in between, the BBQs, movies, and fun in Stockholm! Inge, indrukwekkend wat je allemaal weet en

met jouw humor was het altijd prettig je naast me te hebben! Nikkie, bedankt voor de vrolijkheid, daarmee bracht je zon in de kamer. Jvalini, bedankt voor al de leuke discussies over van alles en voor je betrokkenheid. Neeraj, you give a great balance to the room. Keep on smiling! Ik wens jullie veel succes met het afronden van je proefschrift en de toekomst! I will miss you!

Also my former roommates, Nicole, Natasha, Els and Susan, thanks for all your help and interest in my project.

Narda, de wereld is klein. Wij hebben elkaar voor het eerst ontmoet in Spanje en daarna weer in Wageningen. Dank je voor het eerste contact met dit kleine stadje (wat? Dat was het city center?) en de gezellige tijd!

Die guten, alten ;-) Freunde, Gitte, Sylvia, Uwe, Steven, Verena, ihr habt mich schon 20 Jahre begleitet. Ich will euch von Herzen danken für eure Besuche, und das viele Skypen zwischendurch! Egal wie weit ich weg bin, mit euch zu sprechen ist Heimat... ich hoffe, ihr bleibt eine unverzichtbare Konstante im Leben :-). Gitte, du kannst stolz auf deine kleine Familie sein und wie du nebenbei noch Beruf und Therapeutenausbildung meisterst. Sylvia, ich bewundere dein Durchhaltevermögen, du bist weit gekommen und ich wünsche auch dir mit Marc alles Gute für die Zukunft.

I would like to thank my paranims, Floor and Milène, who were always on my side for the time of my PhD! Floor, although we are both indeed chaotic, I am very thankful that you can stay calm and friendly especially in stressful situations. That made me less worried. Thanks for being my paranim! Milène, zoals jezelf zei, het was wel een avontuur en ik zou het zeker niet gemist willen hebben (onze Sesamstraat). Je hebt een ontzettend steunend en motiverende karakter (op je eigen manier;-). Bedankt voor het vele luisteren! Behoud je optimisme en vrolijkheid! Bedankt ook aan je ouders, Hélène en Richard, voor de steun aan het begin van mijn PhD.

Mutti, Michael, Vati, Oma und Opa, danke für eure Unterstützung, Sorge, alle Freiheit und Interesse in meine Arbeit. Danke für die Pakete gefüllt mit Heimatlekkereien. :-). Ich hab euch alle lieb.

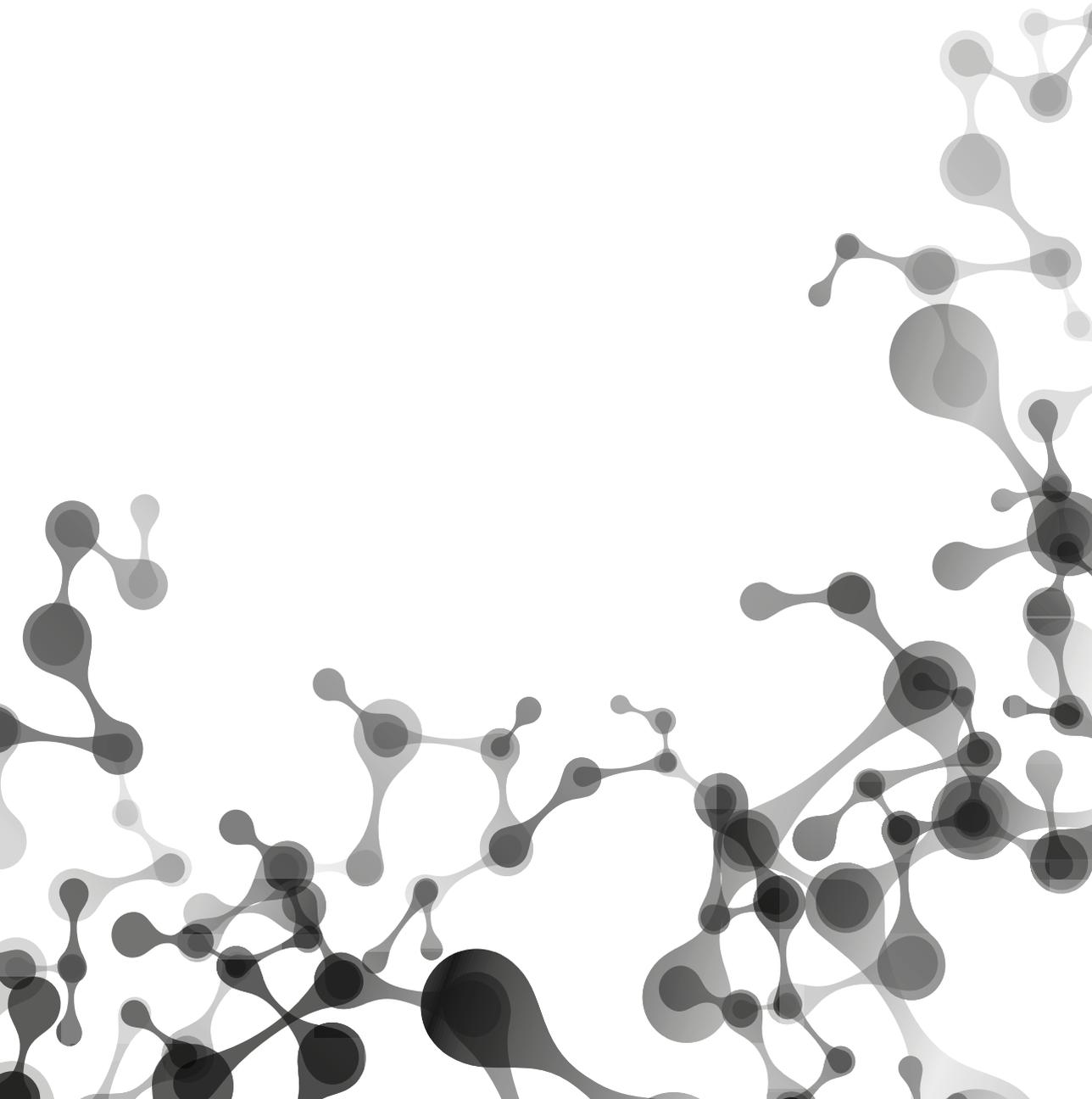
Mijn lieve Matthijs, een glimlach kan heel belangrijk zijn als het mis gaat! Bedankt voor je positiviteit, vrolijkheid, steun en liefde. Je begrijpt me, maakt niet uit in welke taal. Het was de laatste tijd zeker niet makkelijk voor jou en ik ben trots hoe je dat alles voor mekaar hebt gekregen. Ik hoop dat we nog een heel leuke tijd in toekomst samen hebben... ergens ;-). Ik hou van jou!

Danke, Thanks, Bedankt,

Katja



About the author



Curriculum vitae

Katja Lange was born on the 21st November 1984 in Weimar, Germany. She completed secondary school at the Gymnasium in Bad Berka. In 2003, she started her studies in Nutrition Science at the University in Jena, which she finished in 2010. In 2007/2008, Katja studied for one year at the University of Oslo, Norway and worked at the lab of Prof. Andrew Collins conducting experiments using the Comet assay. Her Master thesis she performed at the University of Auckland, New Zealand under supervision of dr. Nishi Karunasinghe and Prof. Lynnette Ferguson. The thesis topic was on selenium, single nucleotide polymorphism and prostate cancer risk. In 2010, Katja started her PhD project at the University of Wageningen at the chair of Nutrition, Metabolism and Genomics, supervised by Prof. Michael Müller and dr. Guido Hooiveld. The results of the PhD thesis 'Molecular mechanisms underlying the effects of dietary fibers in the large intestine' are described in this thesis.

List of publications

This Thesis

Lange K*, Haenen D*, Keshtkar S, Müller M, Hooiveld G. *Transcriptomics revealed distinct short-chain fatty acid-induced changes on colonic gene expression in mice.*

Submitted

Lange K*, Hugenholtz F*, Jonathan M, Schols H, Kleerebezem M, Smidt M, Müller M, Hooiveld G. *Comparison of the effects of five dietary fibers on mucosal transcriptional profiles and luminal microbiota composition and SCFA concentrations in murine colon.*

Submitted

Hugenholtz F*, Lange K*, Davids M, Schaap P, Müller M, Hooiveld G, Kleerebezem M, Smidt H. *Linking the fate of dietary fibers in the murine caecum to microbial transcriptome patterns.*

In preparation

Alex S*, Lange K*, Amolo T, Grinstead JS, Haakonsson AK, Szalowska E, Koppen A, Mudde K, Haenen D, Al-Lahham S, Roelofsen H, Houtman R, van der Burg B, Mandrup S, Bonvin AM, Kalkhoven E, Müller M, Hooiveld GJ, Kersten S. *Short-chain fatty acids stimulate angiotensin-like 4 synthesis in human colon adenocarcinoma cells by activating peroxisome proliferator-activated receptor γ .* Mol Cell Biol, 2013

Lange K, Hugenholtz F*, Ijssennagger N*, Doktorova M, Bortel J, Vervoort J, Mil S, Kleerebezem M, Smidt H, Jonker H, Müller M, Hooiveld G. *Role of epithelial peroxisome proliferator-activated receptor (PPAR) γ in colonic response to fermentation of inulin in mice*

In preparation

*These authors contributed equally to this work

Other:

den Besten G, Lange K, Havinga R, van Dijk TH, Gerding A, van Eunen K, Müller M, Groen AK, Hooiveld GJ, Bakker BM, Reijngoud DJ. *Gut-derived short-chain fatty acids are vividly assimilated into host carbohydrates and lipids.* Am J Physiol Gastrointest Liver Physiol, 2013

Karunasinghe N, Lange K, Yeo Han D, Goudie M, Zhu S, Wang A, Bishop K, Ferguson L, Masters J. *Androgen Pathway Related Gene Variants and Prostate Cancer Association in Auckland Men.* Current Pharmacogenomics and Personalized Medicine, 2013

Karunasinghe N, Han DY, Goudie M, Zhu S, Bishop K, Wang A, Duan H, Lange K, Ko S, Medhora R, Kan ST, Masters J, Ferguson LR. *Prostate disease risk factors among a New Zealand cohort*. J Nutrigenet Nutrigenomics. 2012

Karunasinghe N, Han DY, Zhu S, Yu J, Lange K, Duan H, Medhora R, Singh N, Kan J, Alzahr W, Chen B, Ko S, Triggs CM, Ferguson LR., *Serum selenium and single-nucleotide polymorphisms in genes for selenoproteins: relationship to markers of oxidative stress in men from Auckland, New Zealand*. Genes Nutr, 2012

Overview of completed training activities

Discipline specific activities

System Biology: Statistical Analysis of ~omics, Wageningen, 2010
Netherlands Consortium for Systems Biology (NCSB) Symposium, 2010, 2012, 2013
FEBS-SystemsX.ch Advanced Lecture Course, Innsbruck, 2011
Advanced visualization, integration and biological interpretation of -omics data, Wageningen, 2011
8th NUGO week, Wageningen, 2011
10th NUGO week, Munich, 2013
NUGO masterclass, Wageningen, 2011
5th International Dietary Fibre conference, Rome, 2012
NWO Nutrition Meeting, Deurne, 2010, 2011, 2012, 2013
15th Gut days, Groningen, 2013
Science Camp GRK1482 Technical University (TU) Munich, Seeon, 2013

General courses

Competence Assessment (CA), 2011
Scientific Writing, 2011
Project and Time Management (P&TM), 2011
International Laboratory Animal Science Course, Utrecht, 2011
Master class R statistics - an introduction, Wageningen, 2012
Techniques for Writing and Presenting a Scientific Paper (TWP), 2012
Presentation Skills, 2012

Optional activities

PhD tour 2011 + committee, 2011
TIFN project meetings, 2010-2014
TIFN Theme WE days, 2012
NMG group weekly scientific meetings, 2010-2014
Preparation of research proposal, 2010
Excursion Technical University of Munich, Freising, 2011

Colophon

The research described in this thesis was financially supported by TI Food and Nutrition and Netherlands Consortium for Systems Biology.

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Print

Ipskamp Drukkers, Enschede

Design & lay-out

Promotie In Zicht, Arnhem

© Katja Lange, 2015