Transcriptome analysis of the challenged gut barrier in rats

Mucosal response to Salmonella and Fructo-oligosaccharides

Wendy Rodenburg

Promotor: **Prof. dr. M.B. Katan**

Hoogleraar Voedingsleer

Instituut voor Gezondheidswetenschappen Vrije Universiteit, Amsterdam

Tot 2006: Persoonlijk hoogleraar

Afdeling Humane Voeding, Wageningen Universiteit

Co-promotoren: **Dr. J. Keijer**

Senior wetenschappelijk onderzoeker

Cluster Bioactieve Stoffen RIKILT - Institute of Food Safety, TI Food and Nutrition

Dr. ir. I.M.J. Bovee-Oudenhoven

Projectleider

TI Food and Nutrition, NIZO Food Research

Promotiecommissie: **Prof. dr. M. Müller**

Wageningen Universiteit

Prof. dr. J. Wells

Wageningen Universiteit

Prof. dr. R.J. Brummer

Universiteit Maastricht

Prof. dr. T.A. Niewold

Katholieke Universiteit Leuven, België

Transcriptome analysis of the challenged gut barrier in rats

Mucosal response to Salmonella and Fructo-oligosaccharides

Wendy Rodenburg

Proefschrift

Abstract

Introduction

The gut barrier protects the body against harmful substances and microbes. It consists of the gut mucosa, the immune system and the microflora. Crosstalk between these elements determines the mucosal response to stresses. In this thesis, we have studied how the gut mucosa in a living organism reacts to two stress stimuli: Salmonella and fructo-oligosaccharides (FOS). For this we used transcriptome analysis.

Results

Changes in detoxification, glucose, lipid, peptide and ion transport and proteolysis were part of the early gene expression responses of the rat intestinal mucosa. Neither Salmonella nor FOS altered the expression of barrier related genes such as tight junction, mucin or toll like receptor genes. In contrast, energy metabolism was clearly affected by FOS and could be responsible for the increased permeability induced by FOS. The gene expression response to Salmonella in rats was subtle, this differs from the responses observed in cell culture studies. In contrast to the general expectation, the colon was as much a target for Salmonella as the ileum. FOS increased the expression of Salmonella induced genes, including defence genes, coinciding with increased Salmonella infection. This showed that increase in defence genes reflects a reduced rather than improved gut barrier function, as is often assumed. Several new candidate biomarker genes were identified, such as pancreatitis associated protein (*Pap*), lipocalin, calprotectin and phospholipase A2. PAP protein was studied in more detail. Its response dependent excretion suggests that it can potentially be used as a non invasive marker. Finally, based on the difficulties encountered in analysing transcriptomic data, we propose a framework to identify biologically relevant genes.

Conclusion

We identified biological processes not earlier associated to gut barrier functioning. Expected barrier processes were not induced, thus gut barrier research should not focus on expected barrier processes alone. Extrapolating data from model systems of the barrier to the intact animal should be done with great care, as overlap in gene expression is low. Transcriptome analyses have significantly increased the understanding of the actual *in vivo* barrier processes and have delivered potential new gut health biomarkers.

Content

Chapter 1

General introduction

1. Introduction

2. Intestinal mucosal barrier
2.1 The epithelium
2.2 The immune system
2.3 The microflora

3. External factors affecting the gut barrier
3.1 Intestinal bacterial infections
3.2 Nutritional modulation of intestinal epithelium

4. Transcriptome analysis
4.1 Transcriptomic technique
4.2 From genes to mechanisms
4.3 Intestinal genomics

5. Outline of this thesis

1 Introduction

The primary function of the intestinal epithelium is to digest and absorb nutrients. At the same time it has to prevent infiltration of pathogens and harmful compounds. This barrier function is important for an organism's health and is tightly controlled. Disturbance can be caused by pathogens, harmful compounds or nutrients, and may lead to diarrhoea, infectious disease or uncontrolled inflammation[38]. Reduction of early or mild disturbances could prevent the onset of harmful inflammation reactions. However little is known about the early processes involved in disturbances of the intestinal epithelium.

Understanding of the epithelial barrier responses is limited to model systems such as in vitro cell cultures, or ex vivo epithelial tissue cultures. These models cannot provide a full overview of the multifactorial *in vivo* situation, where the luminal content, the gut microflora and the gut immune system influence the responses of the intestinal epithelium. Therefore, to obtain an overview of the early responses of the intestinal epithelium, examination of the *in vivo* situation is required. Examination of gut barrier mechanisms is possible only after challenging the barrier through induction of stress. This indicates whether the gut barrier is able to resist the stress or not, and it reveals which processes appear to be necessary to resist the stress. In this study, we chose two types of challenges: a pathogenic bacterial challenge and a dietary challenge.

Dietary components can affect the intestinal epithelium directly or indirectly, via changes in the intestinal contents or changes in the endogenous microflora. Human[15] and animal[16,103] studies have shown that diet can modulate intestinal infections. For example, Salmonella infection is sensitive to dietary modulation by calcium and by fructo-oligosaccharides (FOS). Calcium decreases colonization and translocation of Salmonella, whereas FOS increases translocation of this pathogen in rats. However the Salmonella-induced biological processes in the intestinal mucosa and the possible dietary modulation of these processes are not known.

DNA microarray technology allows thousands of genes to be studied at the same time and has been successful in identifying *in vivo* molecular responses of intestinal tissues to commensal bacteria, pathogens or nutrients[111,112]. This technique is not restricted to *a priori* defined biological processes, but identifies all processes active at the time of examination.

The aim of this thesis research was to identify the early gene expression response of the intestinal mucosa in rats to two challenges that adversely affect the barrier function: Salmonella and FOS. We used transcriptomic analysis to look at the whole genome. Increased insight into the molecular response of the gut barrier allows monitoring of gut health and the development of nutrients or pharmaceuticals that are able to modulate early mucosal responses and improve intestinal resistance, for example to infectious disease.

2 Intestinal mucosal barrier

To provide optimal nutrient absorption, the mucosal surface of the intestinal tract is large. In adult humans it represents a surface area of approximately $200m^2$, which is 100 times larger than the surface of the skin. In rats, the gut surface is 25 times larger than the skin surface[25]. This large

surface area, which is in constant contact with the external environment, makes the intestinal mucosa an important target for harmful compounds, such as pathogenic microorganisms, toxins or harmful nutrients. Fortunately, the gastro intestinal (GI) track is equipped with several mechanisms that prevent survival of ingested microorganisms. For example, the acidic environment of the stomach[39], bile[117] and pancreatic enzymes of the small intestine[88], and motility in the small intestine[93]. These mechanisms prevent survival and colonization of microorganisms in the intestinal lumen. Despite these mechanisms, the intestinal epithelium is constantly exposed to unwanted compounds. Successful resistance to these compounds is possible due to the intestinal mucosal barrier. This barrier consists of a monolayer of epithelial cells, the mucosal immune system and the microflora (figure 1). These three components are in continuous interaction with each other.
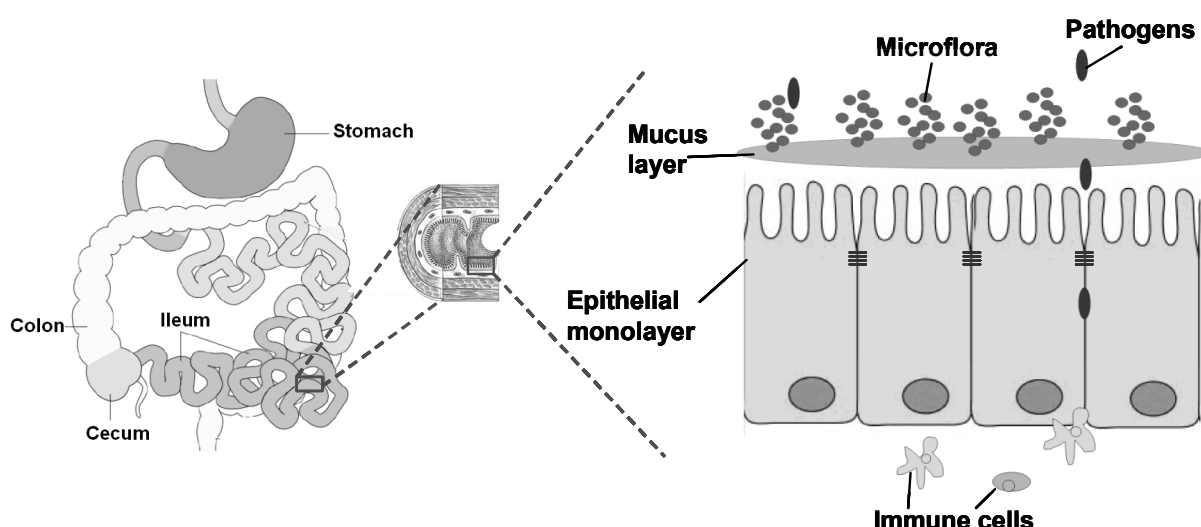

Figure 1. The intestinal barrier aligns the entire gastro-intestinal system.

2.1 The epithelium

The epithelial cells are tightly sealed to each other with tight junctions (TJs). These TJs between the cells regulate intestinal permeability of fluids and small molecules[5,98], and can be affected by pathogens, cytotoxic compounds and cytokines[84,113]. The exact regulation mechanism of TJs is not known, although it is clear that the cytoskeleton is involved[45]. To prevent bacterial invasion through gaps in the epithelial layer, that occur due to loss of damaged cells or apoptosis, adjacent epithelial cells rapidly migrate to the edge of the gap[116]. This process is called regeneration, and after the rapid migration of cells, the epithelium is restored by induced proliferation. Thus regeneration, proliferation and apoptosis are necessary to maintain the epithelial layer and are therefore tightly regulated.

The mucosal epithelial layer is not only a tight monolayer that separates the outside luminal environment from the inside of the body. It also has an important warning function to the immune system, and it actively defends itself against bacteria by producing a sticky mucus layer and antimicrobials[62,74]. The cell types in the intestinal epithelial layer responsible for these functions are enterocytes, goblet cells, Paneth cells, entero-endocrine cells and microfold cells (M-cells).

The most abundant cell type in the intestinal epithelium is the enterocyte. Enterocytes are not only important in nutrient absorption, but also actively communicate with the immune system by producing cytokines and chemokines. In vitro studies have shown that a wide range of stimuli, including pathogens and physical damage, induce the production of cytokines such as interleukin 8, interleuking 1β and tumour necrosis factor α, as well as chemokines such as macrophage inflammatory protein 2 and monocyte chemoattractant protein-1[70,78]. These signalling molecules attract neutrophils and leukocytes to the injured area[100].

Another abundant cell type is the goblet cell. Goblet cells produce a mucus layer that protects the intestinal epithelium against physical damage and infiltration of pathogens simply by acting as a molecular sieve[21,61,72]. Physical, chemical or infectious stimuli trigger goblet cells to release mucin glycoproteins to the luminal side to form the hydrated viscous mucus layer[91]. Goblet cells also produce trefoil factors. These peptides tie mucin glycoproteins together, resulting in increased viscosity and therefore increased protection. Trefoils also play a role in the rapid repair of epithelial damage[58,69]. A wide range of stimuli, including pathogens, dietary fibres or changes in the normal microflora, are able to change mucus composition. The mechanisms responsible for this observed change in composition require further study[24].

Another type of epithelial cell that has a role in the defence are the Paneth cells. These cells are responsible for the production of antimicrobial peptides such as defensins and bacteriolytic enzymes such as trypsin, lyzosyme and phospholipase a2. Antimicrobials are constitutively produced at low levels and are induced by proinflammatory cytokines and exposure to bacteria.

α -Defensin is the most abundant antimicrobial and is able to destroy a wide range of bacteria by disrupting the bacterial cell membrane. Together, these Paneth cell products control the bacterial milieu in the intestine[7,46,73]. In addition, they serve as signalling molecules that communicate between the innate and adaptive immune systems.

2.2 The immune system

The immune system of the intestine has a complex task; it not only has to react to pathogens and harmful compounds, but it also must tolerate the constant flow of food antigens and the abundant non-pathogenic microflora. The intestine is responsible for about 70-80% of the body's immune system, indicating its importance in general defence[35]. The immune system of the intestine contains specialized immune tissues, including the Peyer's patches (PPs) and the mesenteric lymph nodes. Besides these tissues, the immune system consists of several cell types that are scattered through the epithelial mucosal layer, such as dendritic cells, IgA producing B-lymphocytes, and T-lymphocytes[94]. Together, these tissues are called gut-associated lymphoid tissue.

PPs and dendritic cells are essential to the early immune response. The PPs are aggregates of lymphoid tissue in the small intestine that are covered with M-cells[80]. These M-cells constantly sample luminal content to the PPs. The dendritic cells, that are spread throughout the epithelium, sample and transport luminal antigens across the epithelial layer through epithelial TJs[85]. All sampled antigens are presented to B- and T-cells in the PPs, which determine the outcome of an immune response[56]. Upon activation B-cells produce large quantities of immunoglobulin A (IgA) at the mucosal surface. IgA prevents antigens, for instance bacteria, from attaching to mucosal cells and subsequently prevents these bacteria from entering the epithelium. The most important

characteristic of IgA, compared to other immunoglobulins, is that it removes antigens in a non-inflammatory manner[1].

The immune system can be divided into the rapid responding innate immunity and the slow responding adaptive immunity; the rapid innate immune response is non-specific and is active in the first few days of an infection. In addition, it activates the adaptive immune system[97]. This slower responding adaptive response is observed 4 to 7 days after infection and involves the generation of immunological memory. The innate immune system is not only provided by specialized immune cells, but also partly by the intestinal epithelial cells. Both cell types express pattern recognition receptors, such as toll-like receptors, that recognize different pathogen components and play an important role in the host recognition of microflora versus pathogens[86]. For example, flagellin – part of the Salmonella outer membrane – is recognized by toll-like receptor 5 and lipopolysaccharides by toll-like receptor 4. Activation of these various toll-like receptors induce production of a range of cytokines and chemokines by both the enterocytes and immune cells[2,31,79]. Cytokines communicate with the adaptive immune system through either pro-inflammatory cytokines (interleukin 1, tumour necrosis factor α, interferon γ) or anti inflammatory cytokines (interleukin 4, interleukin 10, interferon α). And the chemokines, such as macrophage inflammatory protein 3 α, and mast cell protease 1, that recruit neutrophils, macrophages and dendritic cells to the site of infection[29,30].

2.3 The microflora

The gastrointestinal tract is occupied by a large number of microorganisms, also knows as commensal microflora. In comparison, humans have ten times more bacteria in their gut than cells in their body[12,66]. Most of these bacteria reside in the large intestine. The stomach and small intestine contain relatively few bacteria due to bactericidal activity of gastric acid, bile and pancreatic enzymes, as well as motility in the small intestine[43]. The intestinal microflora bacteria are important for the host because they provides the host with energy and nutrients. These nutrients are produced by bacterial digestion of otherwise indigestible polysaccharides, including plant-derived pectins, cellulose and resistant starches[9]. In addition, the intestinal microflora play an important role in normal gut function and maintaining host health[50]. For example, absence of microflora in germ-free animals results in underdeveloped immune system and underdeveloped metabolism of fatty acids, bile acids and cholesterol[42]. Moreover, the microflora provide a barrier against intruding foodborne pathogens through competition for substrates and mucosal adhesion sites. In addition, organic acid-producing (lactic acid and SCFAs) members of the microflora reduce luminal pH, which inhibits growth of most pathogens[64,109]. The importance of the role for the microflora in intestinal barrier function is indicated by the decreased resistance to luminal pathogens in humans treated with antibiotics[10,102].

The composition of the microflora is influenced by host factors such as antimicrobials and mucus, and external factors such as diet. For instance, almost all carbohydrates that reach the colon provide a substrate for the microflora and affect its growth and metabolic activity. Researchers are very interested in non-digestible oligosaccharides, which are claimed to increase the growth of beneficial bacteria at the expense of less beneficial and pathogenic bacteria; specifically, the lactic acid bacteria bifidobacteria and lactobacilli are increased at the cost of, e.g., bacteroides, clostridia, enterobacteria [13,40].

Together, the epithelium, the immune system and the microflora interact with each other through multiple mechanisms to provide an effective barrier to pathogens (figure 2). Because of these interactions, an *in vivo* approach is required to obtain a complete overview of the intestinal mucosa response to a pathogenic or other stress challenges. However most data on the gut mucosal barrier come from model studies using cell culture studies focusing on one cell type or from *ex vivo* studies where epithelial tissue is studied without luminal content and microflora[51]. These models are important to examine specific issues of barrier defence, but are hard to extrapolate to the *in vivo* situation, where all systems are shown to interact with each other. Furthermore, insight into dietary modulation of intestinal mucosa also requires an *in vivo* approach, since dietary components influence the gut barrier by affecting intestinal content, the composition and activity of the microflora and the functioning of the intestinal epithelium[87].
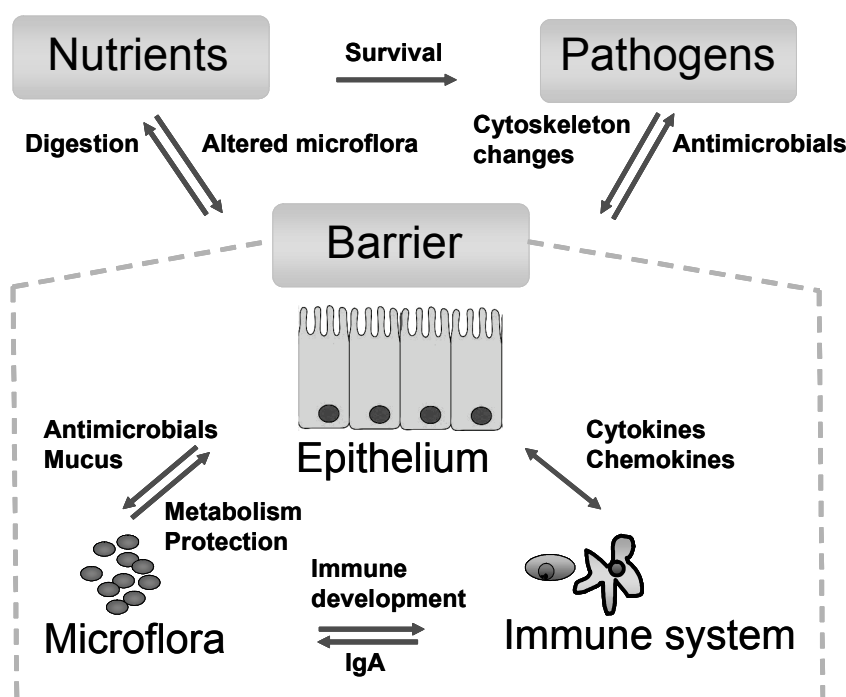
Figure 2. Schematic overview of interactions between the three components that are involved in gut barrier functioning; the epithelium, the immune system and the microflora. External factors such as nutrients and pathogens influence the barrier.

3 External factors affecting the gut barrier

Changes or disruptions of one of the barrier components may signal the onset of intestinal inflammation, resulting in further damage of the intestinal mucosa[27,67]. The early phase of disturbed barrier homeostasis, such as bacterial adhesion or induced intestinal permeability, is sensitive to modulation by nutrients or pharmaceuticals[65,90], for example, dietary calcium was shown to inhibit *Escherichia coli* infection induced diarrhoea[15]. Intervention of this early response could prevent or decrease induction of inflammation. Enhanced mechanistic understanding of the early host mucosal responses to pathogens or harmful compounds is essential for developing dietary products or pharmaceuticals that are able to modulate these early mucosal responses

These early molecular responses of the intestinal mucosa in rats can be examined using challenges that adversely affect the barrier function. Two good model systems in rats are

Salmonella and dietary FOS. Salmonella infection in rats offers a good model system because the infection is sensitive to dietary modulation by calcium and FOS[16,103,105]. Dietary FOS is a good model because FOS increases intestinal permeability and stimulates Salmonella translocation[106], indicating mucosal stress.

3.1 Intestinal bacterial infections

Not all ingested foodborne pathogens induce disease, most pathogens will not survive the upper gastrointestinal-tract defence mechanisms, and the ones that do survive are mainly prevented from adhesion to the intestinal mucosa by the ileal motility and the colonic microflora. However, pathogenic bacteria that are able to adhere to the mucosal surface can start to colonize the gut, and in case of an invasive pathogen they will try to invade the host cells. Bacterial adhesion and invasion may be the most critical events sensitive to dietary modulation and may be a target for strengthening host resistance.

Salmonella

Salmonella is one of the main causes of gastrointestinal infections in Europe and the USA [48,83]. Salmonella is a family of Gram-negative enteropathogenic bacteria that successfully colonizes a wide range of animal hosts, including humans. There are two major groups of Salmonella species: 1) *Salmonella typhi* and *paratyphi,* which mainly infect the small intestine and cause systemic typhoid fever; 2) non-typhoid Salmonella such as *S. typhimurium* and *S. enteritidis*, which induce gastroenteritis with mucosal neutrophil infiltration in the small intestine and the colon. From these species, *Salmonella enteritidis* is the most common cause of food-borne salmonellosis in humans, causing approximately 80% of reported cases in Europe[48] .

Salmonella is frequently used as model pathogen to study fundamental mechanisms of bacterial pathogenesis[82]. However research in living organisms and particularly research on the colon in limited.

Effect of Salmonella on the intestine

In the first step in the infection process, Salmonella adheres to the brush border of intestinal cells[17]. An alternative route of invasion occurs via M-cells covering the Peyer's patches or following capture by dendritic cells[60,80].

Salmonella invades the host by using its type III secretion system, a specialized virulence strategy that injects bacterial proteins into the host cytoplasm. These bacterial proteins influence the host cytoskeleton, resulting in membrane ruffles that enclose the bacteria in large Salmonella-containing vesicles[41,44]. Salmonella modifies these vesicles to prevent their fusion with the lysosomal compartment, enabling survival and bacterial replication[18]. This shows that Salmonella can modulate host cell functions by manipulating the host's cytoskeleton and vesicle machinery. Other host cellular functions that are induced by Salmonella are toll-like receptor 5 signalling and production of antimicrobials[89]. Salmonella-host interactions could affect more processes that have not yet been defined.

Rat model

Early Salmonella-host interactions are most often studied in cell culture systems. Another widely used model in Salmonella research is the murine, bovine or rabbit ligated ileal loop model[51]. This model system is useful to study specific genes or molecules, but lacks both a natural route of infection and the luminal contents important in host defence. A good animal model should

closely mimic the human situation because in the end findings need to be validated in de human situation. The bovine is the model closest to *S. enteritidis* infection in humans. It is not useful for nutritional studies because the gastro-intestinal system of the bovine is completely different than the system in humans. The rat is a good small animal model for Salmonella infection that mimics the human situation[47,76]. For example, the infection of Salmonella is self-limiting in both humans and rats. Where in humans the infection limits to the intestine, in rats systemic infection of the spleen and liver is observed[92]. However, in humans systemic infection is not uncommon in Salmonella-infected elderly, young children and immuno-comprised subjects. Although *in vivo* or *ex vivo* studies on Salmonella infection mainly focus on the ileum, infection of colonic mucosa is also mentioned in humans and rats[68,114]. Therefore both the ileal and the colonic mucosa are important study subjects in Salmonella-host studies.

3.2 Nutritional modulation of the intestinal epithelium

All nutrients pass the intestinal epithelium, and several nutrients are known to ameliorate gastrointestinal integrity. Examples are sulphur-amino acids, glutamine, arginine, zinc, n-3 fatty acids and butyrate[28,119]. There is a great deal of commercial and scientific interest in prebiotics, which are claimed to increase intestinal resistance. However, *in vivo* studies on the role of prebiotics in the defence against foodborne pathogens are scarce[19,20]. In contrast to general expectations, a rat study on the resistance-enhancing properties of the prebiotics FOS showed that dietary FOS increases intestinal permeability and translocation of Salmonella in rats[106]. Furthermore, FOS increased indicators of colonic injury in a rat model for intestinal inflammation[36]. Together these results indicate that FOS induces mild mucosal stress. A diet high in calcium on the other hand, shows to decrease translocation of Salmonella in rats and counteracts the negative effects of FOS on Salmonella translocation[103].

The above results indicate that nutrients can modulate intestinal barrier function. However very little is known about the molecular mechanisms responsible for this dietary modulation of the gut mucosal barrier. Increased insight in these molecular mechanisms could help to develop other nutrients that improve mucosal barrier functioning. FOS has a clear effect on intestinal permeability and on Salmonella infection and is therefore an interesting model nutrient to examine molecular mechanisms involved in gut mucosal barrier functioning.

FOS

FOS is found in many plant species, including wheat, onion, banana and chicory, and is comprised of one terminal glucose molecule and 2-5 fructose units linked by β(2-1) bonds[52]. The digestive enzymes of humans and animals cannot digest the β(2-1) bonds, therefore FOS arrives in the large intestine unchanged where it is hydrolyzed and metabolized by the endogenous microflora. Indeed, when FOS is consumed by ileostoma patients, the average recovery at the terminal ileum lies around 89% of the material consumed[8,34], whereas fecal recovery of FOS in healthy subjects is close to zero[3,71]. The major end products of FOS fermentation by bacteria are lactate and the short chain fatty acids (SCFAs) acetate, propionate and butyrate[115], which are taken up by the epithelial cells. FOS stimulates growth of bifidobacteria and lactobacilli when consumed in amounts of 5-20 grams per day for humans[14,40,107]. Prebiotics are claimed to exclusively promote growth of bifidobacteria and lactobacilli[115], however studies in rats have not

clearly demonstrated this exclusivity[49], in fact FOS has been shown to stimulate the growth of both bifidobacteria and enterobacteria in rats[105].

Effect of FOS on the intestine

As mentioned above, FOS alters the microflora *in vivo*. This can modulate the intestinal barrier function because it has been shown that alterations in the composition of the microfora affect the epithelium and possibly the immune system[50,55]. The fermentation products of FOS, the SCFA's acetate, propionate and butyrate, are also known to affect epithelial cells, as these organic acids provide energy to colonic epithelial cells, decrease luminal pH and influence absorption of calcium and magnesium[23]. Production of modest quantities of SCFAs is essential for normal colonic mucosal function[96]. However, fermentation of large quantities of FOS could cause overproduction of SCFAs. This overproduction may damage the intestinal epithelium, especially when cells are chronically exposed to these high levels of SCFAs[6,63,75]. Thus FOS or its fermentation products clearly affects the mucosal epithelium and its barrier function. Therefore it is interesting to use FOS as a barrier challenge to clarify the molecular mechanisms causing these effects.

Rat model

The rat is a good model organism for nutritional studies on gut health because results from nutritional studies on calcium and FOS on gut health parameters in rats are validated in humans[15,104]. Although the rat and human gastrointestinal tract share the same overall organization, a clear difference is the cecum. In rats the length of the cecum accounts for about 26% of the total length of the large intestine, whereas in humans this is only about 5% [25]. There is also a difference between the quantities of bacteria within the small intestine. In humans, only a few bacteria remain in the small intestine, while in rats 10^6-10^8 CFU per gram content is present[54]. However it has been shown that FOS does not stimulate lactobacilli growth in the rat ileum[106], indicating that FOS fermentation is limited to the distal gut in both rat and human. Furthermore, faecal bifidobacteria, faecal wet weight, faecal lactate and mucin secretion are similarly influenced by FOS in rats and in humans[104].

4 Transcriptome analysis

Transcriptomics, or DNA microarray technology, enables researchers to study thousands of genes at the same time. The technique has rapidly evolved from its introduction in the mid-1990's[95].

A major advantage of this technique is that all biological processes active at the time of examination are identified. New biological processes, formerly not related to the studied subject, can be detected with this technique. This is interesting in studies that examine complex systems, for example the intestine. Previous studies on the *in vivo* response of the intestinal tissue to pharmaceuticals, commensal bacteria, pathogens and nutrients have successfully used transcriptome analysis[50,59,111,112]. Because the molecular mucosal response of the intestine to Salmonella and FOS is an unexplored area of research, this technique is chosen to enable analysis of all biological processes involved in the response.
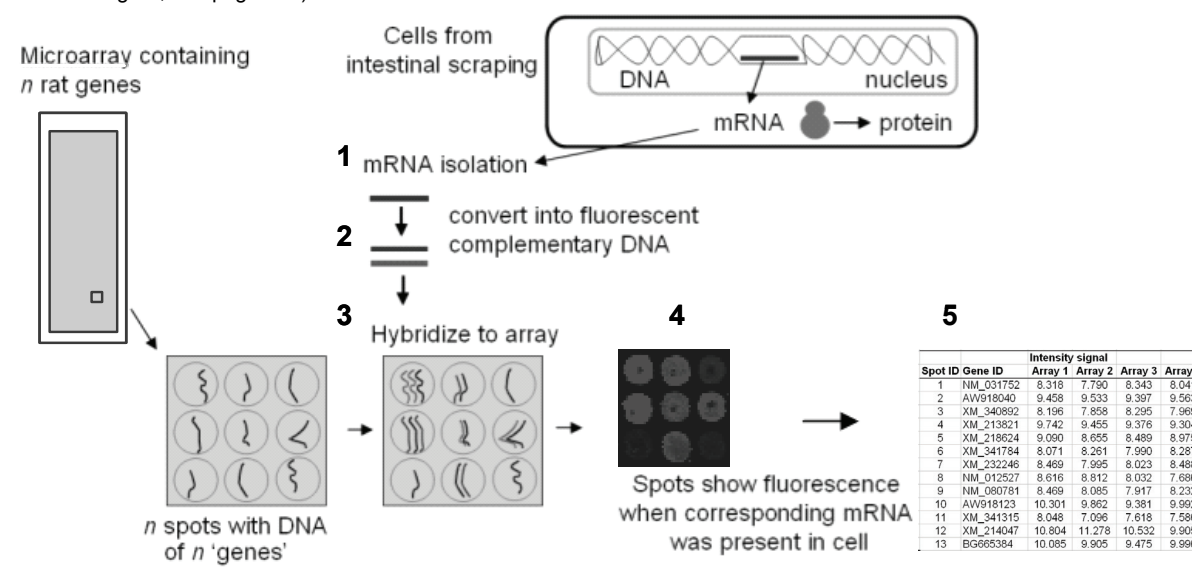
4.1 Transcriptomic technique

Transcriptome analysis identifies and quantifies messenger RNA (mRNA) levels of genes in cells or tissues. The mRNA levels in cells or tissue are indicative of ongoing biological processes. This provides insights into the response of cells or tissues to a wide range of stimuli, including pathogens or nutrients. More traditional methods to examine mRNA levels are Northern blotting[4] and quantitative real-time reverse transcription polymerase chain reaction (Q-PCR)[81]. These methods allow targeted analysis of a relatively small number of genes chosen by the researcher. The microarray technology allows rapid analysis of thousands of genes simultaneously[32], see BOX 1. This type of analysis covers all the biological processes that occur in the cell or tissue under examination. Several types of microarray platforms are available, for example self-spotted cDNA arrays, self-spotted oligo arrays, and commercially available arrays from Agilent[118] or Affymetrix[22].

In addition to large scale study of genes, the proteins and metabolites can also be studied on a large scale; this is called proteomics or metabolomics. Information on proteins and metabolites is closer to physiology than information on gene transcripts. However, transcriptomics generates much larger data sets then proteomics or metabolomics, and there is more data about genes in databases than about proteins and metabolites. In this thesis, the transcriptomic technique is chosen, since this is the most comprehensive technique at this moment.

BOX 1. Transcriptomic technique

The sequences of thousands of genes are available for whole genomes; this allows the design of sequence fragments or probes representing individual genes. These gene-specific probes are printed on nylon membranes, glass slides or silicon chips. Hybridization of single stranded DNA allows determination of mRNA levels of all genes spotted on the array.

Procedure: 1) Isolation of mRNA from biological samples. 2) Fluorescent label incorporation into cDNA copy of the mRNA. 3) Hybridization of labeled sample to the array: each cDNA anneals to the complementary cDNA probe for a specific gene on the surface. 4) Scanning the array to determine the fluorescent amount of labeled cDNA hybridized to each spot.

5) Quantification of the fluorescent signal, and data normalization to enable comparison between the different arrays. (For full colour figure, see page 162)

Spot ID	Gene ID	Intensity signal			
		Array 1	Array 2	Array 3	Array 4
1	NM_031752	8.318	7.790	8.343	8.041
2	AW918040	9.458	9.533	9.397	9.563
3	XM_340892	8.196	7.858	8.295	7.969
4	XM_213821	9.742	9.455	9.376	9.304
5	XM_218624	9.090	8.655	8.489	8.975
6	XM_341784	8.071	8.261	7.990	8.287
7	XM_232246	8.469	7.995	8.023	8.488
8	NM_012527	8.616	8.812	8.032	7.680
9	NM_080781	8.469	8.085	7.917	8.232
10	AW918123	10.301	9.862	9.381	9.992
11	XM_341315	8.048	7.096	7.618	7.580
12	XM_214047	10.804	11.278	10.532	9.905
13	BG665384	10.085	9.905	9.475	9.990

4.2 From genes to mechanism.

The challenging task in transcriptomics is to extract biological data from long lists of gene expression data. As not all genes on the genome are expressed in all tissues, these non-expressed genes are first removed from the dataset. From the remaining gene list, the genes that are differentially expressed under the different experimental conditions need to be identified, because these genes reflect the biological processes changed by the experimental condition. Several approaches exist to obtain biological insight into the list of differentially expressed genes. One approach is to select genes of interest by defining a threshold value. This threshold can be fold change or a test derived p-value, or any other ranking factor derived from the test used. The genes above this threshold are selected for further biological interpretation. The difficulty with these large datasets is the chance of identifying false positive differences between conditions. Statistical tests adapted to transcriptomic data analysis have been developed that can handle these large datasets, such as false discovery rate[11] or significance analysis of microarray data[108], and many other tests are appearing.

Another approach to identify differentially expressed genes is to focus on genes with similar behaviour. Programs such as cluster analysis, self-organizing maps or principle component analysis are able to group genes with similar expression patterns[101,110]. The idea is that genes reacting similarly to a specific stimulus could somehow be functionally related.

All selection options are followed by manual data mining in scientific literature and databases such as Gene Ontology (GO)[37], National Center for Biotechnology Information (NCBI)[77] and Kyoto Encyclopedia of Genes and Genomes (KEGG)[53] to obtain biological information of each selected gene. These approaches are often used, but have the drawback that they are time consuming and depend greatly on the researcher's biological knowledge and preferences.

Biological information of the selected genes can also come from pathway analysis programs, which is more objective and faster than manual data mining. These programs categorize differentially changed genes into biological processes usually based on Gene Ontology categories[57]. In addition to classification in biological processes, these programs can also group genes based on cellular localization or metabolic pathways[33]. The use of pathway programs improves the likelihood of identifying biological processes affected by the experimental treatment and is therefore a requirement in transcriptomic studies[99].

4.3 Intestinal genomics

Altered conditions in the intestinal tract lead to complex interactions between many barrier components, ultimately leading to altered molecular responses in the epithelial cells. Several studies have proven that gene expression studies in epithelial cells can confirm known processes important in gut functioning, but more importantly they all reveal novel insights into gut functioning. Examples are the effect of dietary heme on colonic gene expression[112], altered gene expression in intestinal biopsies from inflammatory bowel disease patients[26], time dependent Trichinella induced gene expression in the mouse intestine[59], and commensal microflora induced gene expression in formerly germfree mice[50]. The newly found processes offer important new understanding and possibilities for further gut research outside the boundaries of common knowledge.

5 Outline of this thesis

The aim of this research was to identify the gene expression response of the intestinal epithelium in rats to two harmful challenges which assault the barrier function; Salmonella and FOS. To obtain insight into the epithelial response to diets and pathogens, an *in vivo* approach is required, because the barrier function of the intestinal mucosa is a result of complex interactions between epithelial cells, luminal content, different epithelial cells and the immune system.
Since little is known about the various intestinal responses that are influenced *in vivo* by Salmonella and FOS, transcriptomics is used to obtain a complete insight into all biological processes that differ between treatment conditions.
This research can help to reveal molecular responses important in maintenance of the intestinal barrier.

Chapter 2 describes the gene expression response of the rat ileum at different times after oral infection of Salmonella. The aim was to identify the point in time when Salmonella induces gene expression changes in the ileum, focusing on both the mucosal layer and the Peyer's patches.

It is widely assumed that Salmonella invades the body through the ileum. However FOS, which affects the colon, has been shown to affect Salmonella infection. **Chapter 3** therefore examines the gene expression of the colon mucosa at different time points after Salmonella infection. It also examines the effect of FOS on Salmonella-induced genes.

In both the ileum and the colon, Salmonella induces Pancreatitis associated protein (PAP) mRNA. The purpose of **Chapter 4** is to examine PAP protein levels in the mucosa and faeces and to define which cells produce PAP in the ileum.

Chapter 5 explores a new method to select subtle gene expression differences from transcriptomic datasets. **Chapter 6** examines the FOS-induced gene expression on colon mucosa without infection to identify which processes are most likely involved in the FOS induced intestinal permeability.

The final Chapter of this thesis (**Chapter 7**) summarizes and discusses the main findings, discusses recommendations for future studies and provides a conclusion.

References

1.   Acheson, D.W. & Luccioli, S. Microbial-gut interactions in health and disease. Mucosal immune responses. *Best practice & research* 18, 387-404 (2004).
2.   Aldridge, P.D., Gray, M.A., Hirst, B.H. & Khan, C.M. Who's talking to whom? Epithelial-bacterial pathogen interactions. *Molecular microbiology* 55, 655-663 (2005).
3.   Alles, M.S*., et al.* Fate of fructo-oligosaccharides in the human intestine. *The British journal of nutrition* 76, 211-221 (1996).
4.   Alwine, J.C., Kemp, D.J. & Stark, G.R. Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. *Proceedings of the National Academy of Sciences of the United States of America* 74, 5350-5354 (1977).
5.   Anderson, J.M., Balda, M.S. & Fanning, A.S. The structure and regulation of tight junctions. *Current opinion in cell biology* 5, 772-778 (1993).
6.   Argenzio, R.A. & Meuten, D.J. Short-chain fatty acids induce reversible injury of porcine colon. *Dig Dis Sci* 36, 1459-1468 (1991).
7.   Ayabe, T*., et al.* Secretion of microbicidal alpha-defensins by intestinal Paneth cells in response to bacteria. *Nature immunology* 1, 113-118 (2000).
8.   Bach Knudsen, K.E. & Hessov, I. Recovery of inulin from Jerusalem artichoke (Helianthus tuberosus L.) in the small intestine of man. *The British journal of nutrition* 74, 101-113 (1995).
9.   Backhed, F., Ley, R.E., Sonnenburg, J.L., Peterson, D.A. & Gordon, J.I. Host-bacterial mutualism in the human intestine. *Science (New York, N.Y* 307, 1915-1920 (2005).
10.   Bartlett, J.G. Antibiotic-associated diarrhea. *Clin Infect Dis* 15, 573-581 (1992).
11.   Benjamini, Y.H., Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc.* 57, 289-300 (1995).
12.   Berg, R.D. The indigenous gastrointestinal microflora. *Trends in microbiology* 4, 430-435 (1996).
13.   Bouhnik, Y*., et al.* Administration of transgalacto-oligosaccharides increases fecal bifidobacteria and modifies colonic fermentation metabolism in healthy humans. *J Nutr* 127, 444-448 (1997).
14.   Bouhnik, Y*., et al.* Effects of fructo-oligosaccharides ingestion on fecal bifidobacteria and selected metabolic indexes of colon carcinogenesis in healthy humans. *Nutrition and cancer* 26, 21-29 (1996).
15.   Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L., Van Doesburg, W., Witteman, B.J. & Van Der Meer, R. Diarrhea caused by enterotoxigenic Escherichia coli infection of humans is inhibited by dietary calcium. *Gastroenterology* 125, 469-476 (2003).
16.   Bovee-Oudenhoven, I.M., Termont, D.S., Weerkamp, A.H., Faassen-Peters, M.A. & Van der Meer, R. Dietary calcium inhibits the intestinal colonization and translocation of Salmonella in rats. *Gastroenterology* 113, 550-557 (1997).
17.   Boyle, E.C. & Finlay, B.B. Bacterial pathogenesis: exploiting cellular adherence. *Current opinion in cell biology* 15, 633-639 (2003).
18.   Brumell, J.H. & Grinstein, S. Salmonella redirects phagosomal maturation. *Current opinion in microbiology* 7, 78-84 (2004).
19.   Buddington, K.K., Donahoo, J.B. & Buddington, R.K. Dietary oligofructose and inulin protect mice from enteric and systemic pathogens and tumor inducers. *J Nutr* 132, 472-477 (2002).
20.   Catala, I*., et al.* Oligofructose contributes to the protective role of bifidobacteria in experimental necrotising enterocolitis in quails. *Journal of medical microbiology* 48, 89-94 (1999).
21.   Corfield, A.P*., et al.* Mucins and mucosal protection in the gastrointestinal tract: new prospects for mucins in the pathology of gastrointestinal disease. *Gut* 47, 589-594 (2000).
22.   Dalma-Weiszhausz, D.D., Warrington, J., Tanimoto, E.Y. & Miyada, C.G. The affymetrix GeneChip platform: an overview. *Methods in enzymology* 410, 3-28 (2006).
23.   Demigne, C., Levrat, M.A., Younes, H. & Remesy, C. Interactions between large intestine fermentation and dietary calcium. *European journal of clinical nutrition* 49 Suppl 3, S235-238 (1995).
24.   Deplancke, B. & Gaskins, H.R. Microbial modulation of innate defense: goblet cells and the intestinal mucus layer. *The American journal of clinical nutrition* 73, 1131S-1141S (2001).
25.   DeSesso, J.M. & Jacobson, C.F. Anatomical and physiological parameters affecting gastrointestinal absorption in humans and rats. *Food Chem Toxicol* 39, 209-228 (2001).
26.   Dieckgraefe, B.K., Stenson, W.F., Korzenik, J.R., Swanson, P.E. & Harrington, C.A. Analysis of mucosal gene expression in inflammatory bowel disease by parallel oligonucleotide arrays. *Physiological genomics* 4, 1-11 (2000).

27.   Ding, L.A. & Li, J.S. Intestinal failure: pathophysiological elements and clinical diseases. *World J Gastroenterol* 10, 930-933 (2004).

28.   Duggan, C., Gannon, J. & Walker, W.A. Protective nutrients and functional foods for the gastrointestinal tract. *The American journal of clinical nutrition* 75, 789-808 (2002).

29.   Dwinell, M.B., Lugering, N., Eckmann, L. & Kagnoff, M.F. Regulated production of interferon-inducible T-cell chemoattractants by human intestinal epithelial cells. *Gastroenterology* 120, 49-59 (2001).

30.   Eckmann, L. & Kagnoff, M.F. Cytokines in host defense against Salmonella. *Microbes and infection / Institut Pasteur* 3, 1191-1200 (2001).

31.   Eckmann, L. & Karin, M. NOD2 and Crohn's disease: loss or gain of function? *Immunity* 22, 661-667 (2005).

32.   Eisen, M.B. & Brown, P.O. DNA arrays for analysis of gene expression. *Methods in enzymology* 303, 179-205 (1999).

33.   Ekins, S., Nikolsky, Y., Bugrim, A., Kirillov, E. & Nikolskaya, T. Pathway mapping tools for analysis of high content data. *Methods in molecular biology (Clifton, N.J* 356, 319-350 (2007).

34.   Ellegard, L., Andersson, H. & Bosaeus, I. Inulin and oligofructose do not influence the absorption of cholesterol, or the excretion of cholesterol, Ca, Mg, Zn, Fe, or bile acids but increases energy excretion in ileostomy subjects. *European journal of clinical nutrition* 51, 1-5 (1997).

35.   Furness, J.B., Kunze, W.A. & Clerc, N. Nutrient tasting and signaling mechanisms in the gut. II. The intestine as a sensory organ: neural, endocrine, and immune responses. *The American journal of physiology* 277, G922-928 (1999).

36.   Geier, M.S., Butler, R.N., Giffard, P.M. & Howarth, G.S. Prebiotic and synbiotic fructooligosaccharide administration fails to reduce the severity of experimental colitis in rats. *Diseases of the colon and rectum* 50, 1061-1069 (2007).

37.   Gene Ontology Consortium. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res* 32, D258-261 (2004).

38.   Gewirtz, A.T., Liu, Y., Sitaraman, S.V. & Madara, J.L. Intestinal epithelial pathobiology: past, present and future. *Best practice & research* 16, 851-867 (2002).

39.   Giannella, R.A., Broitman, S.A. & Zamcheck, N. Influence of gastric acidity on bacterial and parasitic enteric infections. A perspective. *Annals of internal medicine* 78, 271-276 (1973).

40.   Gibson, G.R., Beatty, E.R., Wang, X. & Cummings, J.H. Selective stimulation of bifidobacteria in the human colon by oligofructose and inulin. *Gastroenterology* 108, 975-982 (1995).

41.   Gruenheid, S. & Finlay, B.B. Microbial pathogenesis and cytoskeletal function. *Nature* 422, 775-781 (2003).

42.   Gustafsson, B.E. The physiological importance of the colonic microflora. *Scandinavian journal of gastroenterology* 77, 117-131 (1982).

43.   Hao, W.L. & Lee, Y.K. Microflora of the gastrointestinal tract: a review. *Methods in molecular biology (Clifton, N.J* 268, 491-502 (2004).

44.   Hardt, W.D., Chen, L.M., Schuebel, K.E., Bustelo, X.R. & Galan, J.E. S. typhimurium encodes an activator of Rho GTPases that induces membrane ruffling and nuclear responses in host cells. *Cell* 93, 815-826 (1998).

45.   Hartsock, A. & Nelson, W.J. Adherens and tight junctions: Structure, function and connections to the actin cytoskeleton. *Biochim Biophys Acta* (2007).

46.   Harwig, S.S., Eisenhauer, P.B., Chen, N.P. & Lehrer, R.I. Cryptdins: endogenous antibiotic peptides of small intestinal Paneth cells. *Advances in experimental medicine and biology* 371A, 251-255 (1995).

47.   Havelaar, A.H., *et al.* A rat model for dose-response relationships of Salmonella Enteritidis infection. *Journal of applied microbiology* 91, 442-452 (2001).

48.   Herikstad, H., Motarjemi, Y. & Tauxe, R.V. Salmonella surveillance: a global survey of public health serotyping. *Epidemiology and infection* 129, 1-8 (2002).

49.   Holma, R., Juvonen, P., Asmawi, M.Z., Vapaatalo, H. & Korpela, R. Galacto-oligosaccharides stimulate the growth of bifidobacteria but fail to attenuate inflammation in experimental colitis in rats. *Scand J Gastroenterol* 37, 1042-1047 (2002).

50.   Hooper, L.V., Midtvedt, T. & Gordon, J.I. How host-microbial interactions shape the nutrient environment of the mammalian intestine. *Annual review of nutrition* 22, 283-307 (2002).

51.   Hurley, B.P. & McCormick, B.A. Translating tissue culture results into animal models: the case of Salmonella typhimurium. *Trends in microbiology* 11, 562-569 (2003).

52.   IUB-IUPAC. & Nomenclature., J.c.o.B. Abbreviated terminology of oligosaccharide chains. *J Biol Chem* 257, 3347-3351 (1987).

53. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28, 27-30 (2000).

54. Kararli, T.T. Comparison of the gastrointestinal anatomy, physiology, and biochemistry of humans and commonly used laboratory animals. *Biopharmaceutics & drug disposition* 16, 351-380 (1995).

55. Kelly, D*., et al.* Commensal anaerobic gut bacteria attenuate inflammation by regulating nuclear-cytoplasmic shuttling of PPAR-gamma and RelA. *Nature immunology* 5, 104-112 (2004).

56. Kelsall, B.L. & Rescigno, M. Mucosal dendritic cells in immunity and inflammation. *Nature immunology* 5, 1091-1095 (2004).

57. Khatri, P. & Draghici, S. Ontological analysis of gene expression data: current tools, limitations, and open problems. *Bioinformatics* 21, 3587-3595 (2005).

58. Kindon, H., Pothoulakis, C., Thim, L., Lynch-Devaney, K. & Podolsky, D.K. Trefoil peptide protection of intestinal epithelial barrier function: cooperative interaction with mucin glycoprotein. *Gastroenterology* 109, 516-523 (1995).

59. Knight, P.A*., et al.* Expression profiling reveals novel innate and inflammatory responses in the jejunal epithelial compartment during infection with Trichinella spiralis. *Infection and immunity* 72, 6076-6086 (2004).

60. Kraehenbuhl, J.P. & Neutra, M.R. Epithelial M cells: differentiation and function. *Annual review of cell and developmental biology* 16, 301-332 (2000).

61. Laboisse, C*., et al.* Regulation of mucin exocytosis from intestinal goblet cells. *Biochemical Society transactions* 23, 810-813 (1995).

62. Lievin-Le Moal, V. & Servin, A.L. The front line of enteric host defense against unwelcome intrusion of harmful microorganisms: mucins, antimicrobial peptides, and microbiota. *Clinical microbiology reviews* 19, 315-337 (2006).

63. Lin, J*., et al.* Variable effects of short chain fatty acids and lactic acid in inducing intestinal mucosal injury in newborn rats. *J Pediatr Gastroenterol Nutr* 35, 545-550 (2002).

64. Lopez-Boado, Y.S*., et al.* Bacterial exposure induces and activates matrilysin in mucosal epithelial cells. *The Journal of cell biology* 148, 1305-1315 (2000).

65. Lopez-Varela, S., Gonzalez-Gross, M. & Marcos, A. Functional foods and the immune system: a review. *European journal of clinical nutrition* 56 Suppl 3, S29-33 (2002).

66. Lupp, C. & Finlay, B.B. Intestinal microbiota. *Curr Biol* 15, R235-236 (2005).

67. Macdonald, T.T. & Monteleone, G. Immunity, inflammation, and allergy in the gut. *Science (New York, N.Y* 307, 1920-1925 (2005).

68. Mandal, B.K. & Mani, V. Colonic involvement in salmonellosis. *Lancet* 1, 887-888 (1976).

69. Mashimo, H., Wu, D.C., Podolsky, D.K. & Fishman, M.C. Impaired defense of intestinal mucosa in mice lacking intestinal trefoil factor. *Science (New York, N.Y* 274, 262-265 (1996).

70. Mavris, M. & Sansonetti, P. Microbial-gut interactions in health and disease. Epithelial cell responses. *Best practice & research* 18, 373-386 (2004).

71. Molis, C*., et al.* Digestion, excretion, and energy value of fructooligosaccharides in healthy humans. *The American journal of clinical nutrition* 64, 324-328 (1996).

72. Moncada, D.M., Kammanadiminti, S.J. & Chadee, K. Mucin and Toll-like receptors in host defense against intestinal parasites. *Trends in parasitology* 19, 305-311 (2003).

73. Muller, C.A., Autenrieth, I.B. & Peschel, A. Innate defenses of the intestinal epithelial barrier. *Cell Mol Life Sci* 62, 1297-1307 (2005).

74. Mumy, K.L. & McCormick, B.A. Events at the host-microbial interface of the gastrointestinal tract. II. Role of the intestinal epithelium in pathogen-induced inflammation. *Am J Physiol Gastrointest Liver Physiol* 288, G854-859 (2005).

75. Nafday, S.M*., et al.* Short-chain fatty acids induce colonic mucosal injury in rats with various postnatal ages. *Pediatr Res* 57, 201-204 (2005).

76. Naughton, P.J., Grant, G., Spencer, R.J., Bardocz, S. & Pusztai, A. A rat model of infection by Salmonella typhimurium or Salm. enteritidis. *The Journal of applied bacteriology* 81, 651-656 (1996).

77. NCBI GEO website. http://www.ncbi.nlm.nih.gov/geo/

78. Nemeth, Z.H*., et al.* Disruption of the actin cytoskeleton results in nuclear factor-kappaB activation and inflammatory mediator production in cultured human intestinal epithelial cells. *Journal of cellular physiology* 200, 71-81 (2004).

79. Netea, M.G., van der Graaf, C., Van der Meer, J.W. & Kullberg, B.J. Toll-like receptors and the host defense against microbial pathogens: bringing specificity to the innate-immune system. *Journal of leukocyte biology* 75, 749-755 (2004).

80.	Neutra, M.R., Mantis, N.J. & Kraehenbuhl, J.P. Collaboration of epithelial cells with organized mucosal lymphoid tissues. *Nature immunology* 2, 1004-1009 (2001).

81.	Nolan, T., Hands, R.E. & Bustin, S.A. Quantification of mRNA using real-time RT-PCR. *Nature protocols* 1, 1559-1582 (2006).

82.	Ohl, M.E. & Miller, S.I. Salmonella: a model for bacterial pathogenesis. *Annual review of medicine* 52, 259-274 (2001).

83.	Olsen, S.J*., et al.* The changing epidemiology of Salmonella: trends in serotypes isolated from humans in the United States, 1987-1997. *J Infect Dis* 183, 753-761 (2001).

84.	Pitman, R.S. & Blumberg, R.S. First line of defense: the role of the intestinal epithelium as an active component of the mucosal immune system. *Journal of gastroenterology* 35, 805-814 (2000).

85.	Rescigno, M*., et al.* Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria. *Nature immunology* 2, 361-367 (2001).

86.	Rimoldi, M*., et al.* Intestinal immune homeostasis is regulated by the crosstalk between epithelial cells and dendritic cells. *Nature immunology* 6, 507-514 (2005).

87.	Rowlands, B.J. & Gardiner, K.R. Nutritional modulation of gut inflammation. *The Proceedings of the Nutrition Society* 57, 395-401 (1998).

88.	Rubinstein, E*., et al.* Antibacterial activity of the pancreatic fluid. *Gastroenterology* 88, 927-932 (1985).

89.	Salzman, N.H., Underwood, M.A. & Bevins, C.L. Paneth cells, defensins, and the commensal microbiota: a hypothesis on intimate interplay at the intestinal mucosa. *Seminars in immunology* 19, 70-83 (2007).

90.	Sanderson, I.R. & Naik, S. Dietary regulation of intestinal gene expression. *Annual review of nutrition* 20, 311-338 (2000).

91.	Sansonetti, P.J. War and peace at mucosal surfaces. *Nature reviews* 4, 953-964 (2004).

92.	Santos, R.L*., et al.* Animal models of Salmonella infections: enteritis versus typhoid fever. *Microbes and infection / Institut Pasteur* 3, 1335-1344 (2001).

93.	Sarker, S.A. & Gyr, K. Non-immunological defence mechanisms of the gut. *Gut* 33, 987-993 (1992).

94.	Sato, A. & Iwasaki, A. Peyer's patch dendritic cells as regulators of mucosal adaptive immunity. *Cell Mol Life Sci* 62, 1333-1338 (2005).

95.	Schena, M., Shalon, D., Davis, R.W. & Brown, P.O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science (New York, N.Y* 270, 467-470 (1995).

96.	Scheppach, W. Effects of short chain fatty acids on gut morphology and function. *Gut* 35, S35-38 (1994).

97.	Schnare, M*., et al.* Toll-like receptors control activation of adaptive immune responses. *Nature immunology* 2, 947-950 (2001).

98.	Schneeberger, E.E. & Lynch, R.D. The tight junction: a multifunctional complex. *American journal of physiology* 286, C1213-1228 (2004).

99.	Segal, E., Friedman, N., Kaminski, N., Regev, A. & Koller, D. From signatures to models: understanding cancer using microarrays. *Nature genetics* 37 Suppl, S38-45 (2005).

100.	Stadnyk, A.W. Intestinal epithelial cells as a source of inflammatory cytokines and chemokines. *Canadian journal of gastroenterology = Journal canadien de gastroenterologie* 16, 241-246 (2002).

101.	Tamayo, P*., et al.* Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proceedings of the National Academy of Sciences of the United States of America* 96, 2907-2912 (1999).

102.	Tancrede, C. Role of human microflora in health and disease. *Eur J Clin Microbiol Infect Dis* 11, 1012-1015 (1992).

103.	Ten Bruggencate, S.J., Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L., Katan, M.B. & Van Der Meer, R. Dietary fructo-oligosaccharides and inulin decrease resistance of rats to Salmonella: protective role of calcium. *Gut* 53, 530-535 (2004).

104.	Ten Bruggencate, S.J., Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L., Katan, M.B. & van der Meer, R. Dietary fructooligosaccharides affect intestinal barrier function in healthy men. *J Nutr* 136, 70-74 (2006).

105.	Ten Bruggencate, S.J., Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L. & Van der Meer, R. Dietary fructo-oligosaccharides dose-dependently increase translocation of Salmonella in rats. *J Nutr* 133, 2313-2318 (2003).

106.	Ten Bruggencate, S.J., Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L. & Van der Meer, R. Dietary fructooligosaccharides increase intestinal permeability in rats. *J Nutr* 135, 837-842 (2005).

107.	Tuohy, K.M., Kolida, S., Lustenberger, A.M. & Gibson, G.R. The prebiotic effects of biscuits containing partially hydrolysed guar gum and fructo-oligosaccharides--a human volunteer study. *The British journal of nutrition* 86, 341-348 (2001).

108. Tusher, V.G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences of the United States of America* 98, 5116-5121 (2001).

109. Uribe, A., Alam, M., Johansson, O., Midtvedt, T. & Theodorsson, E. Microflora modulates endocrine cells in the gastrointestinal mucosa of the rat. *Gastroenterology* 107, 1259-1269 (1994).

110. Valafar, F. Pattern recognition techniques in microarray data analysis: a survey. *Ann N Y Acad Sci* 980, 41-64 (2002).

111. van der Meer-van Kraaij, C*., et al.* Differential gene expression in rat colon by dietary heme and calcium. *Carcinogenesis* 26, 73-79 (2005).

112. Van der Meer-Van Kraaij, C., Van Lieshout, E.M., Kramer, E., Van Der Meer, R. & Keijer, J. Mucosal pentraxin (Mptx), a novel rat gene 10-fold down-regulated in colon by dietary heme. *Faseb J* 17, 1277-1285 (2003).

113. Van Itallie, C.M. & Anderson, J.M. The molecular physiology of tight junction pores. *Physiology (Bethesda, Md* 19, 331-338 (2004).

114. Vender, R.J. & Marignani, P. Salmonella colitis presenting as a segmental colitis resembling Crohn's disease. *Dig Dis Sci* 28, 848-851 (1983).

115. Wang, X. & Gibson, G.R. Effects of the in vitro fermentation of oligofructose and inulin by bacteria growing in the human large intestine. *The Journal of applied bacteriology* 75, 373-380 (1993).

116. Watson, A.J*., et al.* Epithelial barrier function in vivo is sustained despite gaps in epithelial layers. *Gastroenterology* 129, 902-912 (2005).

117. Williams, R.C., Showalter, R. & Kern, F., Jr. In vivo effect of bile salts and cholestyramine on intestinal anaerobic bacteria. *Gastroenterology* 69, 483-491 (1975).

118. Wolber, P.K., Collins, P.J., Lucas, A.B., De Witte, A. & Shannon, K.W. The Agilent in situ-synthesized microarray platform. *Methods in enzymology* 410, 28-57 (2006).

119. Ziegler, T.R., Evans, M.E., Fernandez-Estivariz, C. & Jones, D.P. Trophic and cytoprotective nutrition for intestinal adaptation, mucosal repair, and barrier function. *Annual review of nutrition* 23, 229-261 (2003).

Chapter 2

Gene expression response of the rat small intestine following oral Salmonella infection

Wendy Rodenburg[1,2,3]
Ingeborg M.J. Bovee-Oudenhoven[1,3]
Evelien Kramer[1,2]
Roelof van der Meer[1,3]
Jaap Keijer[1,2]

[1] TI Food & Nutrition, Wageningen, The Netherlands
[2] RIKILT Institute of Food Safety, Wageningen, The Netherlands
[3] NIZO Food Research, Ede, The Netherlands

Abstract

Data on the molecular response of the intestine to the food-borne pathogen Salmonella are derived from *in vitro* studies, whereas *in vivo* data are lacking. We performed an oral *S. enteritidis* infection study in Wistar rats to obtain insight in the *in vivo* response in time. Expression profiles of ileal mucosa (IM) and Peyer's patches (PP) were generated using DNA microarrays at days 1, 3, and 6 postinfection. An overview of Salmonella regulated processes was obtained and confirmed by quantitative real-time PCR on pooled and individual samples. Salmonella-induced gene expression responses *in vivo* are fewer and smaller than observed *in vitro* and the response develops over a longer period of time. Few effects are seen at day 1 and mainly occur in IM, suggesting the mucosa as the primary site of invasion. Later, a bigger response is observed, especially in PP. Decreased expression of anti-microbial peptides genes (in IM at day 1) suggests inhibition of this process by Salmonella. Newly identified target processes are carbohydrate transport (increased expression in IM at day 1) and phase I and phase II detoxification (decreased expression at days 3 and 6). Increase of cytokine and chemokine expression occurs at later time points, both in PP and IM. Pancreatitis-associated protein, lipocalin 2, and calprotectin, potential inflammatory marker proteins, showed induced expression from day 3 onward. We conclude that the *in vivo* gene expression response of the ileum to Salmonella differs to a large extent from the response seen *in vitro*.

Introduction

The enteric pathogen *Salmonella enterica* serovar *enteritidis* is one of the main causes of gastro-intestinal infection in Europe and the USA[16,35]. Incidence of infection is highest in children, elderly and immuno-suppressed individuals. In severe cases, illness is the result of translocation from the intestine, mainly the distal ileum, into the bloodstream. Little is known of the response of intestinal cells to the Salmonella infection *in vivo*. Insight in the *in vivo* responses is required to assess the relevance of *in vitro* models and as a starting point to develop strategies to prevent infection, for example by dietary intervention. To obtain an overview of the gene expression response of the ileum mucosa to Salmonella, we used DNA microarray technology. This technology has been used to analyze the gene transcription response of epithelial cell lines exposed to various bacterial pathogens *in vitro*, such as *Vibrio cholerae*, *Listeria monocytogenes*, *Shigella flexneri,* and Salmonella [2,13,36,49]. These pathogens initiated an immune response in the intestinal cell lines within minutes to hours. However, *in vitro* studies do not necessarily represent the *in vivo* situation. The relevant time frame *in vivo* is days rather than hours. Furthermore, monocultures *in vitro* lack interaction with other cell types of the epithelial barrier. *In vitro* cells also lack the microflora and the mucus layer, which are essential components of host-bacterial interaction[19,38,48]. Some of these limitations are overcome in ligated-ileal-loop models, where Salmonella is directly introduced into the ileal lumen. However, these loop models lack luminal contents, and the natural route of pathogenic delivery is omitted. As far as we know, no *in vivo* gene expression studies of the intestine have been performed after oral gavage of Salmonella.

Two distinct tissues in the ileum have been identified as targets for Salmonella entry into the host: the ileum mucosa (IM) and the Peyer's patches (PP) [10,56]. One of the major functions of the IM is to absorb nutrients. Its large surface primarily consists of absorptive enterocytes and secretory Paneth and Goblet cells. PP are immune tissues that are part of the gut-associated lymphoid tissue. M cells are located in the follicle-associated epithelium overlaying the PP that have a role in sampling of pathogens[7,21]. The initial site of Salmonella contact and subsequent translocation in the small intestine is not clear[17,53]. Because the primary site of response to Salmonella is not known in rats we decided to study the course of the molecular response of both target tissues to Salmonella *in vivo*. We chose Wistar rats for our studies, because these have successfully been used in mechanistic infection studies [3,50]. Since diet affects Salmonella colonization and translocation, we used a diet low in calcium and high in fat. This diet results in a low resistance to food-borne bacterial infections[4,5]. As in these studies, we used a dose of Salmonella that causes a self-limiting infection, which is most commonly observed in humans.

To examine time-dependent molecular responses of both IM and PP to oral *S. enteritidis* infection, intestinal tissue of infected and noninfected controls was collected at different time points postinfection (p.i.) (or sham treatment) and pooled samples were analyzed by microarrays. The microarray analysis provided an overview of Salmonella-targeted processes. In our analysis we focused on processes represented by at least three genes showing more than two-fold variation from non-infected controls. A selection of genes from every regulated process was

analyzed by quantitative PCR of RNA samples from individual rats to obtain information about the interindividual biological variation.

Materials and Methods

Animals, diet and infection

The experimental protocol was approved by the animal welfare committee of Wageningen University (Wageningen, the Netherlands). Specific pathogen-free male outbred Wistar rats (WU, Harlan, Horst, the Netherlands) 9 weeks old, mean body weight 285 g, were housed individually in metabolic cages. All animals were kept in a temperature (22-24°C) and humidity controlled (50-60%) room with a 12 h light/dark cycle (lights on from 6 AM to 6 PM). Rats were fed a purified diet during the whole experimental period. Compared with the AIN-93 diet [39], diets were low in calcium (20 mmol CaHPO$_4$·2H$_2$O/kg) and high in fat content (200 g fat/kg)[51], to mimic the composition of a Western human diet. Food and demineralized drinking water were supplied *ad libitum*. Food intake was recorded every day and body weight every 2 days.

Both the control group and the infected group were comprised of 24 rats. Per section time point, eight rats of each group were killed to collect intestinal samples (described below in more detail). The animals were acclimatized to housing and diet for 11 days, after which they were orally infected with *S. enteritidis* (clinical isolate, phage type 4 according to international standards; B1214 culture of NIZO food research, Ede, the Netherlands). In the morning, half of the animals (n= 3 x 8) were orally infected by gastric gavage with 1 ml of saline containing 3 x 10^9 colony forming units (CFU) *S. enteritidis*. The other half of the animals (n = 3 x 8) were sham treated and received saline only. *S. enteritidis* was cultured and stored, as described earlier [5]. Fresh fecal samples were collected on days 1, 2, 3, and 6 p.i. and analyzed for viable Salmonella by plating 10-fold dilutions in sterile saline on modified brilliant green agar (Oxoid, Basingstoke, UK) and incubating aerobically overnight at 37°C. Sulphamandelate (Oxoid) was added to the agar plates to suppress swarming bacteria, such as Proteus species. The detection limit of this method was 10^2 CFU/g fecal wet weight. Total 24 h urine samples were collected on the last day before and on 6 consecutive days after infection. Urine was preserved with oxytetracycline and frozen until analyzed for the nitric oxide metabolites nitrite and nitrate (NOx) by a colorimetric method (Nr. 1746081; Roche diagnostics, Mannheim, Germany). At 1, 3, and 6 days p.i., Salmonella-exposed rats and their corresponding controls were killed by carbon dioxide inhalation. The mesenteric lymph nodes (MLN), spleen, and liver were excised aseptically, weighed, homogenized (Ultraturrax Pro200, Pro Scientific Oxford, CT) in sterile saline, serially diluted, and plated to culture for Salmonella, as described above. The detection limit was 10^2 CFU/g tissue. The distal ileum (defined as the last 12 cm of the small intestine proximal to the cecum) was taken out. The three most distal PP of this intestinal segment were excised and weighed. To obtain IM, the ileum was then longitudinally opened and ileal contents were removed by a quick rinse in cold 154 mM KCl. Subsequently, the mucosa was scraped off using a spatula. The PP and IM were immediately frozen in liquid nitrogen and stored at -80°C for RNA extraction.

RNA isolation

PP and IM scrapings were homogenized in liquid N_2 using a mortar and pestle cooled with liquid N_2 (Fisher Emergo, Landsmeer, The Netherlands). Total RNA was isolated from these homogenates using TRIzol reagent (Invitrogen, San Diego, CA) according to the manufacturer's instructions. Total RNA was purified using RNeasy columns (Qiagen, Chatsworth, CA). Absence of RNA degradation was checked on a 1% Tris-borate-EDTA buffer/agarose gel after 1 hour incubation at 37°C. The purity and concentration were measured with the Nanodrop (Isogen Life Science, Maarssen, The Netherlands). OD A_{260}/A_{280} ratios were all between 2.08 and 2.10, indicating good quality of RNA.

cDNA synthesis

For microarray hybridization, mRNA of eight rats per time point per treatment was pooled. Each pool consisted of equal amounts of RNA of IM or PP from each rat. Arrays were performed in duplicate. For this, RNA pools were split and separately reverse transcribed and labeled with Cy5. A standard reference sample, consisting of a pool of all RNA extracted from IM and PP, was labeled with Cy3. For each oligo array, 40 µg of total RNA was used to make Cy5 or Cy3 labeled cDNA. Total RNA was mixed with 4 µg T21 primer, heated at 65°C for 3 min (RNA denaturation) followed by 25°C for 10 min (primer annealing). cDNA was synthesized by adding 5x first strand buffer (Invitrogen), 10 mM DTT, 0.5 mM dATP, 0.5 mM dGTP, 0.5 mM dTTP, 0.04 mM dCTP, 0.04 mM Cy5-dCTP or Cy3-dCTP, 1.2 U RnaseOUT and 6 U SuperScript II reverse transcriptase to a total volume of 62.5 µL. The reaction was incubated at 42°C for 2 h. Purification, precipitation and denaturation of the labeled cDNA were performed as described earlier [54].

Analysis of mRNA expression by oligo arrays

The rat 10K oligoset (MWG-Biotech, Ebersberg, Germany) used consists of 50-mer rat oligonucleotides representing 9715 rat genes, 100 replicate oligos (8 different genes, 12-14 replicas each), and 169 control oligos (MWG Biotech). The 10K rat MWG oligoset together with an additional set of 104 50-mer oligos representing infection related genes, were printed on Ultra Gaps slides (Corning) using the Microgrid II arrayer (BioRobotics Ltd. Cambridge, UK). After printing, microarrays were allowed to dry at room temperature. The microarrays were immobilized by UV cross linking (120 mJ of UV energy). The microarray slides were prehybridized at 42°C for 4 hours in prehybridization buffer containing 5x SSC, 0.2% SDS, 5x Denhardt's, 200 µg/ml herring sperm DNA, 50% formamide. After prehybridization, all slides were washed twice in MilliQ and once in isopropanol. After being washed, the microarray slides were dried by centrifugation (2 min, 2000rpm). Then the Cy5 labeled cDNAs of the Salmonella infected and control groups were mixed 1:1 with the Cy3 reference labeled cDNA (all in duplicate). Hybridization was performed in a Gene frame (Westburg, the Netherlands) in a volume of 150 µL. The microarrays were hybridized overnight at 42°C in a humid hybridization chamber. After hybridization the arrays were washed and dried as described elsewhere [54]. Arrays were scanned using the Scanarray Express HT (Perkin Elmer, Wellesley, MA) at a laserpower of 90% and a photo-multiplier tube voltage of 55%.

Data analysis

The software package Array Vision (version 7.0, Imaging Research, Ontario, Canada) was used to extract data from the scanned images. Median density values and background values of each spot were extracted for both the (Cy5) and the reference samples (Cy3). Only the spots with average Cy5 and Cy3 values that were twofold above the background value were included in the data analysis. Of the 9819 genes present on the array, 6792 spots fulfilled this criterion. Data normalization was performed with the software package GeneMaths XT (Applied Maths, Sint-Martens-Latem, Belgium) as described elsewhere[37]. The microarray data are deposited in ArrayExpress (http://www.ebi.ac.uk/arrayexpress; E-MEXP-636). The microarray data were analyzed using Microsoft Excel (fold change) and GeneMaths XT (principal component analysis, hierarchical clustering). All groups (controls and infected) were hybridized in duplicate. The few genes with a more than twofold difference between technical duplicates were excluded for further analysis. The noninfected control groups of each time point were hybridized separately. Since no differences between the different days were observed, as analyzed by PCA analysis, the controls were averaged as one group per tissue. Fold changes in transcription levels between Salmonella infected and control samples were calculated from the mean signal values of infected samples of IM or PP on the different time points versus the mean of the control IM or PP. Genes that changed more than twofold at one of the time points compared with either control IM or PP were selected for pathway analysis.

Processes were identified using statistical over-representation in Metacore (GeneGo, St. Joseph, MI), a highly curated web-based application for identification of gene ontology processes in input gene sets[32]. The program uses annotation databases and creates a list of gene ontology processes that are ranked according to their p-value. To assess whether processes were selected by chance, the Metacore pathway analysis was repeated with 15 random sets of 187 genes and the average p-values of each process, of all 15 sets, were used as a surrogate number for false discovery. Since only 30% of the genes were annotated to gene ontology processes, processes with a p-value of <0.01 were manually supplemented with the remaining significant genes using biological databases (BIOCarta, Gene Ontology, GenMAPP, KEGG) and scientific literature. Genes with analogous function or unequivocally being part of the same functional process or pathway were included, whereas far-away members were excluded. To conclude that a biological process was differentially affected in infected versus control, two criteria were used: 1) initial p-value in Metacore had to be smaller than p<0.01 and 2) at least three genes of that process had to be changed more than twofold upon Salmonella infection. The use of two criteria for selection was used to prevent overinterpretation and thus possible misinterpretation.

Analysis of mRNA expression by quantitative real-time PCR

Quantitative real-time PCR (Q-PCR) on individual samples and pooled samples was performed to confirm differences in mRNA levels. It was considered unnecessary to analyze all (3 x 8) non-infected control animals individually because the array data of the control groups killed at days 1, 3, and 6 were highly similar. Instead, nine control animals were randomly chosen for individual RT verification. One microgram of RNA of all individual samples was used for the cDNA synthesis using the iScript cDNA synthesis kit of Bio-Rad Laboratories (Veenendaal, The

Netherlands). Real-time reactions were performed by means of the iQ SYBR Green Supermix of Bio-Rad using the MyIQ single-color real-time PCR detection system (Bio-Rad). Each reaction (25 μl) contained 12.5 μl of iQ SYBR green supermix, 1 μl of forward primer (400 nM), 1 μl of reverse primer (400 nM), 8.5 μl of RNase-free water and 2 μl of diluted cDNA. The following cycles were performed 1 x 3 min at 95°C, 40 amplification cycles (40 x 10 s 95°C, 45 s 60°C), 1 x 1 min at 95 °C, 1 x 1 min at 62 °C and a melting curve (80 x 10 s at 55 °C with an increase of 0.5 °C per 10 s). A negative control without cDNA template was run with every assay. The optimal melting point of dsDNA (Tm) and the efficiency of the reaction were optimized beforehand. A Tm of 60 °C was chosen for all reactions, and a PCR efficiency of 90-110% (3.2 < slope > 3.8) together with a correlation coefficient of > 0.99 were accepted. Data were normalized against the reference genes β-actin (*Actb*) and pleckstrin homology domain-containing family A member 3 (*Plekha3*). Primers were designed using Beacon designer 4 (Premier Biosoft International, Palo Alto, CA). For sequences, see supplemental table 2.1. A standard curve for all genes, including reference genes was generated using serial dilutions of a pooled sample (cDNA from all reactions). mRNA levels were determined from the appropriate standard curve. Samples with mRNA levels below the lowest standard value were given a value not lower than half the value of this lowest standard, corresponding to the detection level. Analysis of all individual samples was performed in duplicate. Statistical analysis of the data was performed in Prism 4 (Prism 4, GraphPad software, San Diego, CA) using Student's *t*-test. $p<0.05$ (*) was considered statistically significant and $p<0.01$ (**) highly significant.

Results

Physiological response to Salmonella
In agreement with previous studies, food consumption and growth of the Wistar rats were not affected by Salmonella infection. High translocation of Salmonella to mesenteric lymph nodes, but not the liver and spleen, was seen at day 1 (table 1). This implies that at day 1, Salmonella has already crossed the intestinal barrier. Spleen and liver showed colonization by Salmonella at later time points. NO_x, as a parameter of systemic infection, was found to be increased from day 3 onward (table 1).

Gene expression induced by Salmonella in IM and PP
The oligonucleotide micorarray contained 9715 genes. After hybridization of RNA from ileal mucosa (IM) and Peyer's patches (PP), the expression of 6792 genes was more than twofold above the background in one or both tissues. The expression of 187 genes (98 genes in IM, 128 genes in PP) was altered more than two fold in at least one of the Salmonella-infected groups compared with the uninfected controls (days 1, 3, and 6) (Supplemental table 2.3 and Supplemental table 2.4 are available at the website of Physiological Genomics).

Table 1. Viable Salmonella counts in feces, mesenteric lymph nodes, spleen, and liver and NO_x concentrations in urine of rats 1, 3 and 6 days post infection.

		Salmonella (logCFU/g)[a]		
	Control	Day 1 p.i.	Day 3 p.i.	Day 6 p.i.
Feces	N.D.	7.22±0.19	5.92±0.24	6.04±0.32
MLN	N.D.	3.38±0.43	5.85±0.13	5.44±0.05
Spleen	N.D.	N.D	3.20±0.33	3.49±0.05
Liver	N.D.	N.D	N.D	2.45±0.12
NO_x (μmol/day)[b]	6.56±0.23	6.84±0.22	9.62±1.13	33.44±5.13

[a] The rats were orally infected with 3 x 10^9 colony-forming units of *S. enteritidis* or sham treated. Salmonella counts are expressed in log values as means ± SEM (n=8). N.D. = not detected.
[b] NO_x values are means   SEM (n=8).

(http://physiolgenomics.physiology.org/). For subsequent analyses, the non-infected controls of the three different section days were taken as one group, since their expression pattern fully overlapped. No differentially expressed genes (cut-off ratio of >1.5) could be identified when the three control groups were compared to each other (data not shown). At day 1 p.i., only small effects in Salmonella-induced gene expression were observed. Just five genes were affected more than two fold in the PP. A larger response was observed in IM, where 18 genes showed differential expression compared to controls. At days 3 and 6 p.i., a larger response was seen in both tissues and, in contrast to day 1, the response in PP was stronger than the response in IM. Most affected genes showed increased expression by Salmonella during the course of infection. However, a notable portion of the genes in IM at day 6 p.i. showed decreased expression (figure 1). Some overlap in Salmonella-induced genes was seen in both tissues (figure 2). Assessment of epithelial, goblet and Paneth cell specific genes (data not shown) and several well known housekeeping genes (see table 3) indicated that the observed gene expression data did not result from changes in cellular composition of the mucosa.

To further characterize tissue-specific and common responses, the set of 187 genes showing at least twofold increased or decreased expression compared with the noninfected control level were classified in biological processes. Processes with p<0.01 were inferred to be meaningfully related to the Salmonella response. Several randomly selected sets of 187 genes were also classified into processes, for these sets the significance for all selected processes, obtained by Metacore pathway analysis, was not significant (p>0.05). Clearly, the significance of the processes identified using the 187 Salmonella affected genes was much higher, making identification by chance highly unlikely.

To prevent the occurrence of false positive genes, and overinterpretation of biological processes affected by Salmonella, we focused on biological processes with at least three genes exceeding the cut-off >2.0. Additionally we observed that the genes within all functional groups, except for lipid and other transporters, showed a comparable pattern of expression (table 3). Also, most processes showed a similar pattern of expression in both IM as well as in PP, which strongly indicates that these processes are truly affected by Salmonella.

A

B

Figure 1. The number of differentially expressed genes with a fold change greater than 2 in ileal mucosa and ileal Peyer's patches of rats orally infected with Salmonella at 1, 3 or 6 days postinfection compared to sham treated controls (A).

Hierarchical clustering of genes with a fold change greater than 2 (infected/control). Genes and experiments were clustered using Pearson UPGMA (GeneMathsXT). Color scale represents ratio infected/control (B). (For full color figure, see page 163).

The majority of the differentially affected genes could be grouped into the following processes: immune response, inflammation, antimicrobial defense, complement cascade, detoxification, transport, and extracellular matrix organization (table 2). Genes belonging to these processes, but with a differential expression of 1.5 to 2.0- fold, are also included in the tables (tables 2 and 3).

To confirm the Salmonella-induced effect on biological processes selected, we performed an independent array hybridization of freshly pooled ileal mucosal RNA from the same infection study. The results of this analysis confirmed all selected genes based on the two selection criteria applied; 1) twofold change cut-off; 2) at least three genes changed in a similar biological process. For the present study, this corroborates that the used selection criteria were robust and valid.

| | **Mucosa** | | | **Peyer's patches** | |
| | total | only | shared | only | total |

Day 1　　18　　　13　　5　　　　　　　5

Day 3　　45　　　21　　24　　66　　　90

Day 6　　58　　46　　12　　66　　　78

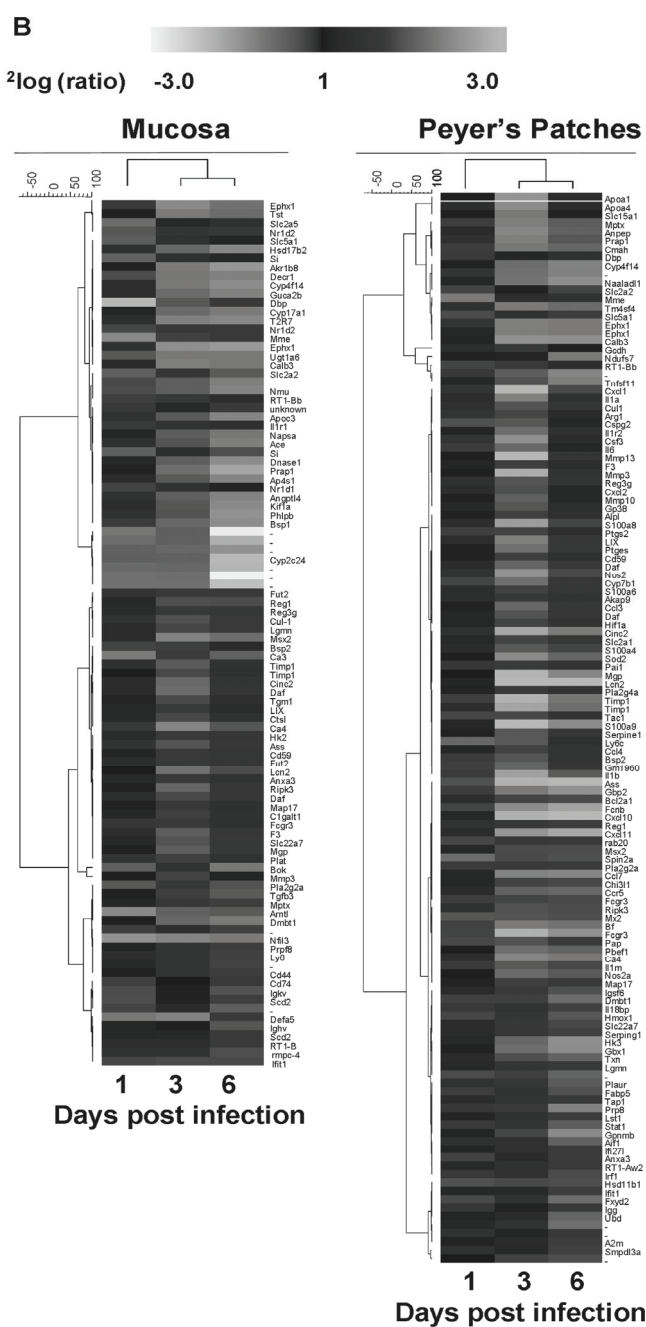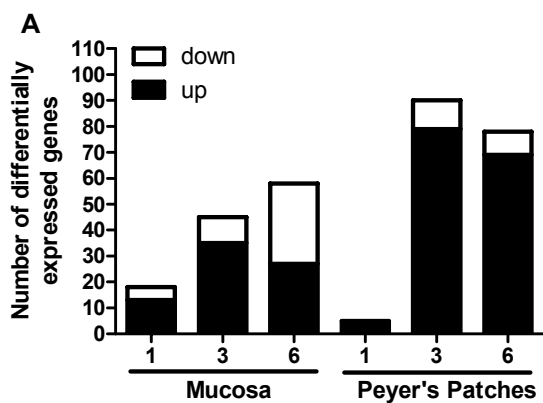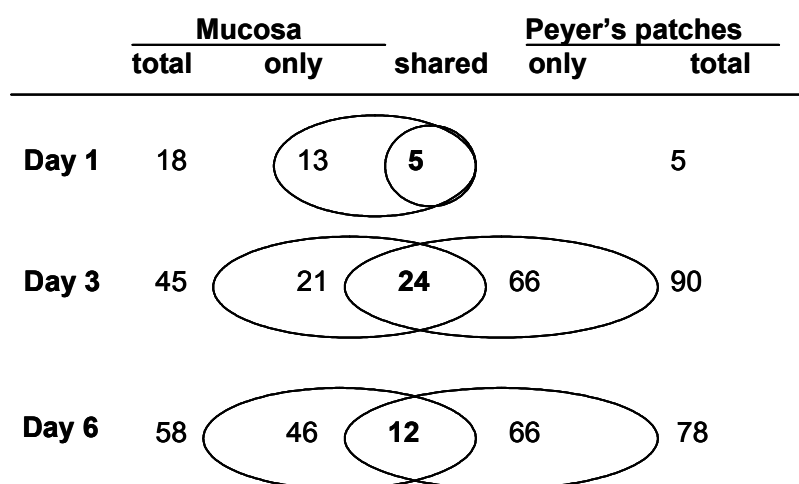**Figure 2.** Comparative transcription profiles of differentially expressed genes with a fold change greater than 2 in ileal mucosa and ileal Peyer's patches of rats orally infected with Salmonella at 1, 3 or 6 days postinfection compared to sham treated controls. For both tissues the total, unique, and shared genes are presented. The numbers of shared genes modulated by Salmonella are indicted in bold type.

Immune activation, inflammation and antimicrobial defense

Genes encoding antimicrobial defense proteins defensin 5 (*Def5*), lysozyme (*Lys*), and matrilysin (*Mmp7*) showed a 1.7 to 2.0-fold decreased expression in IM at day 1 p.i. (table 3). Salmonella did not stimulate the expression of genes related to the innate immune response and inflammation at day 1 p.i. The decreased expression of antimicrobial defense genes had mostly disappeared at days 3 and 6 p.i., whereas expression of those related to immune response and inflammation was clearly induced, with the strongest response observed at day 3. The PP showed a more pronounced immune and inflammatory response than the IM at days 3 and 6 p.i. (table 3). Lipocalin 2 (*Lcn2*) and pancreatitis-associated protein 3 (*Pap3*) encode for inflammatory response proteins whose expression was affected in PP and IM. Furthermore, one activator of the complement cascade, tissue factor (coagulation factor) III, and three inhibitors of this cascade, Decay-accelerating factor (*Daf*), *Cd59* (protectin), and serpinG (C1 inhibitor), showed induced expression on Salmonella infection in both tissues (table3).

Table 2. Biological processes in ileal mucosa and Peyer's patches affected by Salmonella infection.

	Process	Mucosa[a]		Peyer's patches[a]	
		up	down	up	down
Day 1					
	Carbohydrate transport	3	-	3	-
	Transcription	2	2	1	-
	Antimicrobial	-	2	-	-
Day 3					
	Immune response	7	-	32	-
	Detoxification	-	3	1	-
	Extra cellular matrix	3	-	7	-
	Complement	3	-	4	-
	Transcription	2	1	-	-
	Inflammation	-	-	7	-
Day 6					
	Immune response	9	-	35	-
	Detoxification	-	6	1	3
	Inflammation	-	-	6	-
	Extra cellular matrix	-	-	3	-

[a] Numbers represent the amount of genes within that process exceeding a twofold or greater differential expression between oral infected and sham-treated rats at day 1, 3 and 6 postinfection.

Table 3. Gene expression changes in genes related to immune activation, inflammation, antimicrobial defense, complement pathway, detoxification, transport and housekeeping.

Gene Name[a]	Gene symbol	Common alternative	Accession number	Fold Change[b]					
				Mucosa			Peyer's patches		
				1	3	6	1	3	6
Activation immune system									
Interleukin 1 beta	*Il-1β*	*Il1b*	NM_031512	-	1.9	-	-	**6.0**	**2.8**
Interleukin 1 alpha	*Il-1a*	*Il1a*	NM_017019	-	1.5	-	-	**4.0**	1.5
Interleukin 6	*Il-6*	*Il6*	NM_012589	-	-	-	-	**3.2**	-
Chemokine (C-X-C motif) ligand 1	*CXCL1*	*Groα*	NM_030845	-	1.6	-	-	**11.4**	**2.4**
Chemokine (C-X-C motif) ligand 2	*MIP-2*	*Groβ*	NM_053647	-	-	-	-	**2.5**	-
Neutrophil chemoattractant-2 beta	*MIP-2B*	*Groγ*	NM_138522	-	**3.2**	1.7	-	**6.4**	**3.8**
Chemokine (C-X-C motif) ligand 5	*Cinc-5*	*Lix*	NM_022214	-	**2.1**	-	-	**4.0**	1.9
Chemokine (C-X-Cmotif) ligand 10	*Cinc-10*	*IP10*	U22520	-	1.8	1.6	-	**7.3**	**9.5**
Chemokine (C-X-Cmotif) ligand 11	*Cinc-11*	*I-TAC*	AF179872	-	-	-	-	**5.2**	**6.6**
Chemokine (C-C motif) ligand 3	*CCL3*	*Mip-1A*	NM_013025	-	-	-	-	**3.2**	1.9
Chemokine (C-C motif) ligand 4	*CCL4*	*Mip-1B*	U06434	-	-	-	-	**2.2**	1.7
immunoglobulin alpha-chain	*IgA*	-	M13801	-1.6	-1.7	1.7	-	-	-
immunoglobulin gamma-2a	*IgG2a*	-	L22654	-	-	-	-	-	**2.0**
immunoglobulin gamma-2b	*IgG2b*	-	M13802	-	-	1.7	-	-1.6	1.7
Inflammation markers									
Cyclooxygenase-2	*Cox-2*	*Ptgs2*	NM_017232	-	-	-	-	**2.2**	-
Nitric oxide synthase	*Nos2*	*iNOS*	NM_012611	-	1.7	-	-	**5.0**	**2.4**
Superoxide dismutase	*Sod2*	*MnSod*	NM_017051	-	-	-	-	**4.5**	**3.1**
Calgranulin A	*S100a8*	Calpro	NM_053822	-	-	-	-	**10.2**	**4.6**
Calgranulin B	*S100a9*	-tectin	NM_053587	-	1.8	-	-	**5.8**	**2.3**
Lipocalin 2	*Lcn2*	*Ngal*	X13295	-	**3.5**	**2.4**	-	**34.3**	**12.7**
Pancreatitis associated protein III	*Reg3γ*	*Pap3*	U09193	-	**2.0**	1.9	-	**2.6**	1.7
Antimicrobial defense, Paneth cell derived									
Defensin 5	*Def5*	*Rd5*	AF115768	**-2.0**	**-2.2**	-	-	-	-
Lysozym	*Lys*	-	NM_012771	**-2.0**	-1.8	-	-	-	-
Matrilysin	*MMP7*	-	NM_012864	-1.7	-	1.7	-	-	-
Phospholipase a2, group iia	*Pla2g2a*	-	NM_031598	-1.4	1.9	**2.6**	-	**2.0**	**2.1**
Complement pathway									
Tissue factor f3	*F3*	*CD142*	NM_013057	-	**2.9**	1.8	-	**2.1**	-
Serine (or cysteine) peptidase inhibitor, clade G, member 1	*Serping1*	*C1Inh*	NM_199093	-	1.9	-	-	1.8	**2.2**
Decay accelerating factor GPI-form	*CD55*	*Daf*	AB026903	-	**2.8**	1.8	-	**2.8**	1.9
CD59 glycoprotein	*CD59*	*Protectin*	NM_012925	-	**2.0**	1.7	-	**2.1**	-
Detoxification Phase I									
Cytochrome p450 1a1	*Cyp1a1*	-	NM_012540	-	-1.8	-	-	-	-
Cytochrome p450 2j4	*Cyp2j4*	-	NM_023025	-	-	**-2.0**	-	-	-
Cytochrome p450 2c24	*Cyp2c24*	-	S59652	-	-1.6	**-3.8**	-	-	-1.6
Cytochrome p450 3a9	*Cyp3a9*	-	U60085	-	-1.5	-1.8	-	-1.6	-1.5
Cytochrome p450 4f1	*Cyp4f1*	-	NM_019623	-	**-2.0**	**-2.4**	-	-1.8	**-2.2**
Cytochrome p450 17a1	*Cyp17a1*	-	NM_012753	-	-1.8	**-2.1**	-	-1.7	-1.9
Cytochrome p450 7b1	*Cyp7b1*	-	U36992	-	1.8	-	-	**3.4**	1.9
Epoxide hydrolase 1	*Ephx1*	-	NM_012844	-	**-2.4**	**-2.2**	-	-1.7	-1.9
Carboxylesterase 1	*Ces1*	-	NM_031565	-	-1.7	-1.7	-	-1.8	-1.7
Detoxification Phase II									
UDP glucuronosyltransferase 1	*Ugt1a6*	-	D38067	-	**-2.0**	**-2.0**	-1.6	-	-1.8
Glutathione S-transferase alpha 2	*Gsta2*	-	NM_017013	-	-1.7	-	-	-1.8	-

Glucose transporters									
glucose transporter Glut 5	Slc2a5	Glut5	NM_031741	**3.3**	-	1.5	1.8	-	-
na+/glucose cotransporter SGLT1	Slc5a1	Sglt1	NM_013033	**2.9**	1.5	-	**2.4**	-	-
glucose transporter Glut 2	Slc2a2	Glut2	NM_012879	**2.9**	1.7	-	**2.3**	-	-
glucose-transporter protein	Slc2a1	Glut1	M22063	-	-	-	-	1.9	1.7
Lipid transporters									
apolipoprotein b	Apob	-	U53873	1.9	-	-	1.9	-	-
Apolipoprotein A-I	Apoa1	-	NM_012738	-	-	-	-	**-2.6**	-
Apolipoprotein A-IV	Apoa4	-	NM_012737	-	-1.7	-1.7	-	**-2.4**	-
lipid-binding protein	Fabp5	E-Fabp	U13253	-	**2.0**	1.7	-	-	**2.5**
sodium-dependent bile acid transporter	Slc10a2	-	NM_017222	1.9	-	1.5	-	-	-
Other transporters									
organic anion transporter Phase I	Slc22a7	Oat2	NM_053537	-	**2.6**	1.9	-	1.9	**2.3**
oligopeptide transporter, member 1	Slc15a1	Pept1	NM_057121	1.8	-	-	-	**-2.0**	-
Housekeeping genes[c]									
Actin beta	Actb	-	NM_031144	-1.1	1.2	1.2	-1.4	1.5	-1.1
Glyceraldehyde-3-phosphate dehydrogenase	Gapdh	-	NM_017008	1.1	1.1	-1.1	1.2	1.1	1.4
Aldolase a, fructose-bisphosphate	Aldoa	-	NM_012495	-1.1	-1.1	1.1	-1.1	1.3	1.4
Phosphoglycerate kinase 1	Pgk	-	NM_053291	1.1	1.3	1.0	1.2	1.3	1.3
Adp-ribosylation factor 1	Arf1	-	NM_022518	1.2	-1.2	1.1	-1.2	1.0	1.1
H3 histone, family 3B	H3f3b	-	BC086580	-1.1	1.0	1.0	-1.1	-1.1	1.0
Succinate dehydrogenase complex, subunit A	Sdha	-	NP_569112	1.0	-1.1	1.0	1.1	-1.1	-1.1
Pleckstrin homology domain-containing family A member 3	Plekha3	-	NM_001013077	1.0	1.0	1.0	1.0	1.1	1.0

[a] Genes with a fold-change in expression between rats that were orally infected by Salmonella relative to their sham-treated controls. The genes with fold-change greater than 1.5 are given for day 1, 3 and 6 postinfection in ileal mucosa and Peyer's patches.

[b] Fold change > 2.0 is shown in bold; a minus sign indicates lower expression in Salmonella-infected animals.

[c] The fold change in expression of housekeeping genes is given for all groups.

Detoxification genes

A second group of affected genes were detoxification genes, which showed lower expression at days 3 and 6 p.i. (table 3). This group consisted of both phase I and II genes. Phase I genes encoding cytochrome P450 1a1 (*Cyp1a1*) , *Cyp2j4*, *Cyp2c24*, *Cyp3a9*, *Cyp4f1*, *Cyp17a1*, epoxide hydroxylase 1 (*Ephx1*), and carboxylesterase 1 (*Ces1*) showed lower expression, with the exception of *Cyp7b1* expression which was higher. Expression of two phase II genes, encoding UDP-glucuronosyltransferase 1 (*Ugt1*) and Glutathione S-transferase alpha 2 (*Gsta2*), was also lower.

Transporters

A third category of genes affected by Salmonella infection was transporters (table 3). Three carbohydrate transporters, the apical located *Sglt1* and *Glut5* and basolateral located *Glut2* showed higher expression, primarily at day 1 in both tissues studied. The other transporters that were affected showed diverse expression patterns (table 3).

Validation of Salmonella-regulated genes in individual animals by Q-PCR.

To verify the Salmonella modulation of certain biological processes, we selected genes from every process for Q-PCR confirmation. Pooled Q-PCR analysis was performed on genes involved in antimicrobial defense (*Def 5*, *Lys*, *Mmp7*), chemotaxis (*Groa*), inflammation (*Lcn2, Pap3*), phase I detoxification metabolism (*Cyp3a9, Cyp4f1, Cyp1a1, Ephx*), and glucose transport (*Slc5a1, Slc2a2, Slc2a5*). For individual analysis the genes *Def5, Lcn2, Pap3, Cyp4f1* and

Table 4. Gene expression differences (fold change) analyzed by DNA microarray and Q-PCR.

	Gene symbol	Accession number	Fold Change[a].					
			Ileal Mucosa			Peyer's patches		
			Micro array	Q-PCR pool[b]	Q-PCR individual[c]	Microarray	Q-PCR pool[b]	Q-PCR individual[c]
Day 1	*Def5*	AF115768	-2.0	-	-2.2	-	-	-
	Lys	NM_012771	-2.0	-3.3	-	-	-	-
	MMP7	NM_012864	-1.7	-3.3	-	-	-	-
	Slc2a5	NM_031741	3.3	3.2	-	1.8	4.0	-
	Slc2a2	NM_012879	2.9	3.8	-	2.3	8.4	-
	Slc5a1	NM_013033	2.9	1.9	2.5	2.4	3.5	5.2
Day 3	*Groa*	NM_030845	1.6	-	3.0	11.4	-	13.0
	Pap 3	U09193	2.0	-	4.2	2.6	-	3.7
	Lcn	X13295	3.5	-	5.2	34.3	-	30.1
	Cyp1a1	NM_012540	-1.8	-5.0	-	-	-	-
	Cyp3a9	U60085	-1.5	-2.0	-	-1.6	-1.4	-
	Cyp4f1	NM_019623	-2.0	-2.5	-1.6	-1.8	-2.0	-2.0
	Ephx	NM_012844	-2.4	-2.5	-	-1.7	-2.0	-
Day6	*Cyp3a9*	U60085	-1.8	-2.5	-	-	-	-
	Cyp4f1	NM_019623	-2.4	-5.0	-3.6	-2.2	-2.5	-2.0
	Ephx	NM_012844	-2.2	-2.5	-	-1.9	-1.3	-

[a] Data are given as fold-change between rats that were orally infected by Salmonella and their sham-treated controls.

[b] Fold change based on Q-PCR data of pooled samples.

[c] Average fold change based on Q-PCR data of individual samples (n=8). The fold-change for Q-PCR was normalized against reference gene β-actin.

Slc5a1 were chosen. β-Actin (*Actb*) and pleckstrin homology domain-containing family A member 3 (*Plekha3*) were selected as reference genes because they showed constant and treatment-independent expression in the array data (data not shown). All Q-PCR analyses on individual and pooled samples of both IM and PP confirmed the microarray data since relative gene expression changes were similar using both methods (see table 4). In noninfected control animals, the expression of *Groa* and *Lcn2* was close to background, which may have affected the precise fold change. Analysis of the individual samples by Q-PCR generally revealed large inter-individual variation in gene expression within treatment groups (figure 3). The gene expression of several genes known to be affected by Salmonella in *in vitro* studies such as *Tlr5* and the nuclear factor κb (Nf-κb) subunits *RelA* and *P105* were not found to be affected based on array data. Q-PCR analysis confirmed the lack of induction in this *in vivo* study (figure 3).
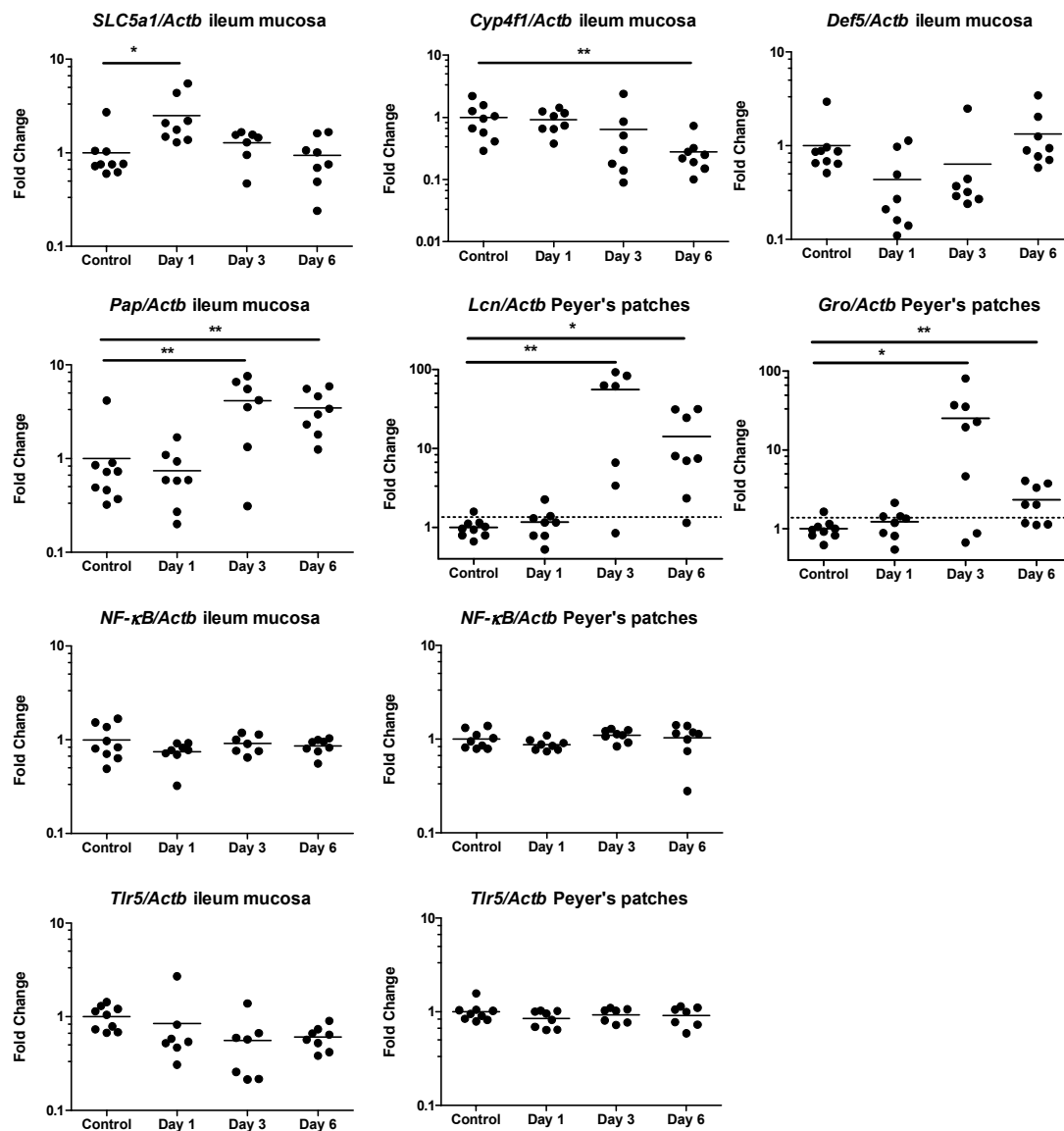
Figure 3. Individual mRNA expression of Na+/glucose cotransporter (*Slc5a1*), cytochrome P450 4f1 (*Cyp4f1*), defensin 5 (*Def5*), pancreatitis-associated protein 3 (*Pap3*), lipocalin 2 (*Lcn2*), chemokine *Groα*, Nuclear factor kappa-B, p105 (*NF-κB*) and Toll-like receptor 5 (*Tlr5*) of sham treated control rats (mean is set as 1) and of rats at 1, 3 and 6 days after oral infection by Salmonella as analyzed by Q-PCR. The expressions of the genes are relative to β-actin in ileal mucosa and in Peyer's patches for the individual animals. Dotted line indicate lowest mRNA standard. The mean gene expression value within each group is indicated ▬. The asterisk indicates p< 0.05(*), *p*<0.01 (**), when comparing the infected group and the control group. *Def5* normalized with reference gene β-actin gave *p*= 0.07 for day 1 p.i., normalization with *Plekha3* reached *p*<0.05.

Discussion

Overall *in vivo* response

In this study, we used microarrays to follow the *Salmonella enteritidis*-induced gene expression changes in the ileum of Wistar rats with time. In two target tissues, the IM and the ileal PP, only a very limited number of genes was changed at 1 day p.i. Altogether, only 0.2% - 0.9% of the genes in IM show a differential gene expression at day 1 through day 6 p.i. This contrasts with *in*

vitro responses of human intestinal epithelial cells, where 5-35% of the genes were found to be affected within the first 20 hours after infection with Salmonella compared with noninfected controls[13]. Infection of intestinal cell lines (T84, CaCo-2) with other enteric bacteria for up to 3 hours resulted in 4% differentially expressed genes compared to non-infected cells [36,49]. Moreover, not only the number but also the magnitude of most responses in our *in vivo* study are relatively small compared with gene expression changes *in vitro*, which may change up to 90-fold. Studies in bovine ileal loops infected with *S. typhimurium* showed an eight-fold increased expression of pro-inflammatory chemokine and cytokine genes within 6 hour p.i.[43]. In fact, no correlation (r^2=0.0004) was found if the genes that were found to be more than twofold altered *in vitro* in HT29 cells 3 hours after infection with Salmonella [13] were compared with the *in vivo* response of the same 30 genes that were present in this study (see supplemental figure 2.1, and supplemental table 2.2).

In vitro models can provide insight in mechanistic aspects of Salmonella-host interaction *in vivo*. However, the large gene expression responses observed in Salmonella infection studies with HT-29 cells[13] are in contrast with the limited gene expression response observed in Salmonella infected intact mucosa or PP presented in our *in vivo* study. The overlap in genes observed to change in *in vitro* studies[13] and our *in vivo* study is very small. Several chemokines (*Mip2/ Mip2a, Groa/Il-8*) are induced in both type of studies, but even then the time frame (3 hours *in vitro* versus 3 days *in vivo*) and magnitude clearly differ (91.5 and 78.0 fold *in vitro* and 1.4- and 2.5-fold *in vivo* for *Il-8* and *Mip2a*, respectively).

The *in vitro* response to Salmonella infection seems unphysiologically reactive, possibly due to the absence of a protective intestinal microflora, mucus layer, and mucosal secretion of antimicrobials. As a result of the absence of these barriers, the number of invasive Salmonella's per cell is likely much higher *in vitro* than *in vivo*, and this will affect the response of these cells.

Additionally, whereas *in vitro* homogeneous cell types are studied, a natural heterogeneous mixture of cell types is studied *in vivo*. This may affect the type and magnitude of the gene expression response. It should be realized that a response of a limited number of specific cells to Salmonella at early time points of the infection might have been missed in the present study due to possible dilution of these cells in the heterogeneous cell populations present in the ileal mucosal scrapings and PP. Identification of cell type-specific responses of, e.g., dendritic cells or other potential target cells can be addressed using laser microdissection to isolate a specific cell type before RNA isolation.

The absence of gene expression effects of expected chemokines and cytokines at day 1 in the present *in vivo* study and the larger gene expression effects *in vitro* and *ex vivo* are in line with differences in phenotypic, physiological observations. Where *in vitro* systems show massive cell death at 24 hours and ileal loop models show epithelial detachment after 8 hours of exposure to Salmonella[63], no inflammatory changes are seen at the first day after oral infection with Salmonella *in vivo*[44]. This late *in vivo* response is also seen in our study, where NO_x a marker of the a-specific immune response, increased from day 3 onward. It should be noted that at day 1 day p.i. high Salmonella numbers were observed in the MLNs in the present and previous studies[4,5], demonstrating translocation of Salmonella at this early time point.

That translocation of this pathogen did not provoke extensive early gene expression changes *in vivo* may indicate that Salmonella infection is a targeted and controlled process. In living

organisms gut epithelial cells are in continuous and intimate contact with gut bacteria. It is known that these host-microbe interactions are important for keeping inflammatory processes in check. The inflammatory response can be repressed by the microflora,[9,19,24], Also, the surrounding host cells suppress signals, e.g. epithelial cell-derived factors influence dendritic cell responses which may regulate the generation of an inflammatory response to bacteria[41]. Communication and feedback mechanisms between different mucosal cell types help to maintain mucosal homeostasis. Cell lines *in vitro* miss contact with other cell types and the modulating effect that commensal bacteria may have.

Another possible, but less likely, explanation for the small gene expression changes observed in this *in vivo* study could be related to host specificity of Salmonella-induced responses. The pathogenicity of Salmonella serovars can be animal species-specific[18,46,55] and Salmonella *in vitro* studies are mostly performed in human cell lines[20]. However, many aspects, such as time-course effects of Salmonella colonization and translocation, are largely similar in this rat model compared with humans [3,5,15].

Site of Salmonella invasion, mucosal glucose metabolism and defense

Based on the number of genes altered by Salmonella, the IM showed an earlier response than the PP. This may indicate that in rat the IM, and not PP, is the first site of interaction or invasion. Within this early response, we newly identified that Salmonella changed genes related to glucose metabolism. These glucose metabolism-related genes, sodium-dependent glucose transporter (*Sglt1*), fructose transporters *Glut5* and *Glut2*, and sucrase-isomaltase (*SI*) (fold change 2.9 in IM at day 1 p.i.), are expressed in enterocytes, which implies that the enterocytes are the first contact or entry site of Salmonella infection. The expression of *Sglt1*, *Glut,5* and *Glut2* in the PP most likely originates from enterocytes overlying the PP[45]. The upregulated glucose transport may be triggered by a higher glucose need of infected cells, which has been reported in cells infected by chlamydia and viruses[34,47]. Altered cell metabolism may also explain the differential expression of other transporters (e.g. lipid transports *Apob, Apoa1* and *4, E-Fabp, SLC10a2*, peptide transporter *Pept-1* and organic anion transporter *Oat2*). Salmonella decreased the expression of genes important for host defense against bacterial intruders [(defensin 5 (*Def5*), lysozyme (*Lys*), matrilysin (*Mmp7*) and secretory phospholipase A_2 (*Pla2g2a*)] at the early time point. Reduced expression of α-defensin and lysozyme was also reported in mice inoculated with *S. typhimurium*[42]. Defensin, lysozyme and matrilysin are expressed by Paneth cells located in the bottom of the mucosal crypts. These early changes suggest that, among the different cell types lining the IM (enterocytes, Paneth cells, goblet cells, and neuro-endocrine cells), the Paneth cells and enterocytes are a target for Salmonella. Based on these results it seems that the IM is at least as important as the PP as the major site of early Salmonella invasion. This enterocyte-targeted invasion of Salmonella has also been reported in experiments with calves [14], pigs [27], and rabbits [57].

Detoxification

Phase I and phase II detoxification enzymes in both IM and PP showed decreased expression, coinciding with increased expression of inflammatory genes. This was not caused by dilution of epithelial cells, since the expression of *I-Fabp*, a control for epithelial content[31], showed no significant differential expression. One explanation for the decreased expression of detoxification

genes might be that this allows optimal defense by immune cells. Down-regulation of cytochrome P450 gene expression is known to be induced by inflammatory mediators such as reactive oxygen species, nitric oxide, IFN, or cytokines (Il-1, transforming growth factor β)[29,40] and is also observed in mucosal biopsies from IBD patients [26]. Alternatively, the expression may be actively reduced by Salmonella, especially since the Ah receptor nuclear translocator (*Arntl*), which is the transcription factor regulating *Cyp1a1, Ugt1a6* and *Gsta2*[61], shows decreased expression at day 1 (fold change -2,1 in IM).

Innate immune response

Despite clear translocation of Salmonella from the intestinal lumen to the MLN (table 1), no increased expression of immune response genes was seen at day 1 p.i., neither in the PP nor in the IM (tables 2 and 3). So far, most of the studies focusing on host gene expression responses on exposure to micro-organisms have been performed *in vitro* [20]. At later time points, we observe *in vivo* some genes (*Il-1a, Il-1β, Groa, Groγ, iNos*) that are a confirmation of former *in vitro* infection studies[20], but we also observe that some genes well known to be upregulated by Salmonella in *in vitro* studies[62], such as *Tlr5* and the Nf-κb subunits *RelA* and *P105*, were not found to be affected based on array data. To exclude that the absence of differential gene expression was due to a technical issue, we examined the differential expression of *Tlr5* by Q-PCR analysis in individual samples and confirmed the absence of differential expression (figure 3). Also the main downstream signaling molecule, *NF-κB p105* subunit, was not regulated (figure 3). This was also the case for the NF-κB *relA* subunit (data not shown).

Despite the absence of gene expression changes, this pathway seems to be activated, since we observed differential expression of targets of TLR and NF-κB activation such as cytokine, chemokine and inflammatory response genes (*Il-6, Groa, β, γ, iNos, Cox2, Sod2*)[22,52], at the later time points. That the immune response genes could not be observed at day 1 may result from induction in a limited number of cells or specific cell types, which are diluted in the heterogeneous cell population of the mucosal scrapings and PP. Another explanation might be modulation at posttranscription level, which escapes detection at transcription level.

Recruitment

Salmonella invasion is characterized by recruitment of monocytes, neutrophils and dendritic cells to the infected area[60]. Indeed, later time-points showed a prominent increase in the expression of genes involved in chemotaxis, including several CC-chemokines (*CCL3, CCL4*) and several CXC chemokines (*CXCL1, MIP-2, MIP-2B, Cinc-5, Cinc-10, Cinc-11*). These genes were among the most highly differentially expressed genes in this study, particularly in the PP. Another prominent group of genes in PP are inflammation related genes. The stronger induction of genes involved in chemotaxis and inflammation in PP compared with IM suggests either a higher Salmonella invasion of the PP at later time points or a stronger secondary response induced by more recruitment of leukocytes to this tissue.

Damage control

Protective mechanisms against epithelial barrier disruption show differential expression at later time points. This late induction is most likely stimulated to limit inflammation-induced damage to

the mucosa. This is best reflected in the increased expression of *Daf* and *Cd59*. These are inhibitors of the complement cascade and aid restoration of blood flow in the microvasculature.

Markers

At days 3 and 6 on Salmonella infection, expression of the general inflammatory mediators *Cox2, iNOS* and *Sod2, Pap3, Lcn* and calprotectin (*S100A8* and *S100A9*) was strongly increased. Calprotectin, *Pap3* and *Lcn,* are all highly expressed in the chronically inflamed mucosa of inflammatory bowel disease (IBD) patients and in animal models of this disease[6,12,23,33,58]. Obviously, generic mechanisms are involved in acute inflammation due to Salmonella infection and chronic inflammation in IBD, despite different pathologies. Possibly, *Pap, Lcn2* and calprotectin could serve as general markers for gastrointestinal inflammation. Since these markers are all secreted into stool, resistant to degradation by intestinal contents, and easily measured [6,11,30], they might be very useful to follow the course of an inflammatory period by non-invasive means.

Responses driven by Salmonella

The present study focused on Salmonella-induced changes in intestinal processes. Some identified processes such as defense and immune response are a confirmation of known effects of pathogens on the host. But others, such as changes in detoxification and transporter genes, are not related to infection before. We cannot discriminate whether these induced changes are directly caused by Salmonella or secondary effects of infection-induced inflammation. Induction by Salmonella itself cannot be excluded, since it is known that microbes are able to induce host gene expression for their own benefit. Salmonella uses its own type III secretion system to alter host cell processes, including apoptosis, cytoskeletal function and cytokine production[1,25,28]. Other than pathogens, commensals can also actively induce host gene expression and thus affect important physiological functions[8,9]. Most of these results originate from *in vitro* studies. The *in vivo* relevance for (intestinal) host resistance and gut barrier functioning should be addressed in future studies.

Technical aspects

Microarray analysis provided an overview of processes in rat ileum that are affected by oral infection of Salmonella. The processes that were obtained were not selected using random sets of 187 genes (data not shown), indicating that these processes are truly affected by Salmonella and not selected by chance. To further confirm the selection of processes, selected genes, representative of various physiological processes, were investigated by Q-PCR in individual rats. The individual genes showed a statistically significant change within the group (n = 8 rats), indicating that these genes are truly affected by Salmonella and not purely by chance. These analyses revealed a high variation in expression among different rats within a treatment group, a finding most likely resulting from the genetic heterogeneity of the outbred Wistar rats used in the present study. Use of inbred animals likely reduces heterogeneity but has the disadvantage that observed effects may be specific for a particular genetic background[59] and thus hamper translation to humans. The large interanimal variation observed has implications for future studies. Using the same setup, it will be difficult to identify differential responses below twofold.

Such a twofold differential response may already constitute a relevant and large physiological response, especially if this occurs in several genes in the same pathway simultaneously. Despite the relevance of relatively small changes in gene expression, it will be technically difficult to study, e.g., the preventative effects of dietary components or the therapeutic effects of new drugs on infection-induced processes by current PCR- or array-based methods. To deal with this limitation and to confidently identify smaller effects, it would be best to include more animals per treatment group and to perform array hybridization of individual samples. This allows identification of differentially expressed genes, not only based on magnitude of change but also on statistical power.

Conclusion

In conclusion, the present *in vivo* study reveals that *Salmonella enteritidis* induces small gene expression changes in the ileum of Wistar rats. Especially at day 1 p.i. a very limited response in gene expression was observed despite marked translocation of Salmonella to the MLNs. Remarkably, the few early changes observed occurred in the IM. This may indicate that IM but not PP is the primary target for Salmonella translocation, at least in rats. The more various gene expression changes at days 3 and 6 p.i. were mainly observed in the PP and were related to immune cell recruitment and inflammation. Infection-induced inflammatory genes overlap with those reported to be up-regulated in inflammatory bowel disease, and may thus be explored as general markers of intestinal inflammation. Finally, we newly identified that mucosal glucose metabolism and detoxification capacity are affected by Salmonella infection in the rat.

Acknowledgements

References

1.   Abrahams, G.L. & Hensel, M. Manipulating cellular transport and immune responses: dynamic interactions between intracellular Salmonella enterica and its host cells. *Cell Microbiol* 8, 728-737 (2006).

2.   Baldwin, D.N., Vanchinathan, V., Brown, P.O. & Theriot, J.A. A gene-expression program reflecting the innate immune response of cultured intestinal epithelial cells to infection by Listeria monocytogenes. *Genome Biol* 4, R2 (2003).

3.   Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L., Van Doesburg, W., Witteman, B.J. & Van Der Meer, R. Diarrhea caused by enterotoxigenic Escherichia coli infection of humans is inhibited by dietary calcium. *Gastroenterology* 125, 469-476 (2003).

4.   Bovee-Oudenhoven, I.M., ten Bruggencate, S.J., Lettink-Wissink, M.L. & van der Meer, R. Dietary fructo-oligosaccharides and lactulose inhibit intestinal colonisation but stimulate translocation of Salmonella in rats. *Gut* 52, 1572-1578 (2003).

5.   Bovee-Oudenhoven, I.M., Termont, D.S., Weerkamp, A.H., Faassen-Peters, M.A. & Van der Meer, R. Dietary calcium inhibits the intestinal colonization and translocation of Salmonella in rats. *Gastroenterology* 113, 550-557 (1997).

6.   Carlson, M.*, et al.* Human neutrophil lipocalin is a unique marker of neutrophil inflammation in ulcerative colitis and proctitis. *Gut* 50, 501-506 (2002).

7.   Clark, M.A. & Jepson, M.A. Intestinal M cells and their role in bacterial infection. *Int J Med Microbiol* 293, 17-39 (2003).

8.   Clarke, M.B. & Sperandio, V. Events at the host-microbial interface of the gastrointestinal tract III. Cell-to-cell signaling among microbial flora, host, and pathogens: there is a whole lot of talking going on. *Am J Physiol Gastrointest Liver Physiol* 288, G1105-1109 (2005).

9.   Collier-Hyams, L.S., Sloane, V., Batten, B.C. & Neish, A.S. Cutting edge: bacterial modulation of epithelial signaling via changes in neddylation of cullin-1. *J Immunol* 175, 4194-4198 (2005).

10.   Darwin, K.H. & Miller, V.L. Molecular basis of the interaction of Salmonella with the intestinal mucosa. *Clin Microbiol Rev* 12, 405-428 (1999).

11.   Desjeux, A.*, et al.* Serum measurements of pancreatitis associated protein in active Crohn's disease with ileal location. *Gastroenterol Clin Biol* 26, 23-28 (2002).

12.   Dieckgraefe, B.K.*, et al.* Expression of the regenerating gene family in inflammatory bowel disease mucosa: Reg Ialpha upregulation, processing, and antiapoptotic activity. *J Investig Med* 50, 421-434 (2002).

13.   Eckmann, L., Smith, J.R., Housley, M.P., Dwinell, M.B. & Kagnoff, M.F. Analysis by high density cDNA arrays of altered gene expression in human intestinal epithelial cells in response to infection with the invasive enteric bacteria Salmonella. *J Biol Chem* 275, 14084-14094 (2000).

14.   Frost, A.J., Bland, A.P. & Wallis, T.S. The early dynamic response of the calf ileal epithelium to Salmonella typhimurium. *Vet Pathol* 34, 369-386 (1997).

15.   Havelaar, A.H.*, et al.* A rat model for dose-response relationships of Salmonella Enteritidis infection. *J Appl Microbiol* 91, 442-452 (2001).

16.   Herikstad, H., Motarjemi, Y. & Tauxe, R.V. Salmonella surveillance: a global survey of public health serotyping. *Epidemiol Infect* 129, 1-8 (2002).

17.   Hughes, E.A. & Galan, J.E. Immune response to Salmonella: location, location, location? *Immunity* 16, 325-328 (2002).

18.   Ishibashi, Y. & Arai, T. A possible mechanism for host-specific pathogenesis of Salmonella serovars. *Microb Pathog* 21, 435-446 (1996).

19.   Ismail, A.S. & Hooper, L.V. Epithelial cells and their neighbors. IV. Bacterial contributions to intestinal epithelial barrier integrity. *Am J Physiol Gastrointest Liver Physiol* 289, G779-784 (2005).

20.   Jenner, R.G. & Young, R.A. Insights into host responses against pathogens from transcriptional profiling. *Nat Rev Microbiol* 3, 281-294 (2005).

21.   Jepson, M.A. & Clark, M.A. The role of M cells in Salmonella infection. *Microbes Infect* 3, 1183-1190 (2001).

22.   Jiang, B.*, et al.* Temporal control of NF-kappaB activation by ERK differentially regulates interleukin-1beta-induced gene expression. *J Biol Chem* 279, 1323-1329 (2004).

23.   Keilbaugh, S.A.*, et al.* Activation of RegIIIbeta/gamma and interferon gamma expression in the intestinal tract of SCID mice: an innate response to bacterial colonisation of the gut. *Gut* 54, 623-629 (2005).

24.   Kelly, D.*, et al.* Commensal anaerobic gut bacteria attenuate inflammation by regulating nuclear-cytoplasmic shuttling of PPAR-gamma and RelA. *Nat Immunol* 5, 104-112 (2004).

25.   Kuhle, V., Abrahams, G.L. & Hensel, M. Intracellular Salmonella enterica redirect exocytic transport processes in a Salmonella pathogenicity island 2-dependent manner. *Traffic* 7, 716-730 (2006).

26.   Langmann, T.*, et al.* Loss of detoxification in inflammatory bowel disease: dysregulation of pregnane X receptor target genes. *Gastroenterology* 127, 26-40 (2004).

27.   Meyerholz, D.K.*, et al.* Early epithelial invasion by Salmonella enterica serovar Typhimurium DT104 in the swine ileum. *Vet Pathol* 39, 712-720 (2002).

28.   Miller, S.I. & Pegues, D.A. Salmonella species, including Salmonella typhi. in *Principles and practice of infectious disease* (ed. Mandell, G.L.) 1700-1716 (2000).

29.   Morgan, E.T., Li-Masters, T. & Cheng, P.Y. Mechanisms of cytochrome P450 regulation by inflammatory mediators. *Toxicology* 181-182, 207-210 (2002).

30.   Naughton, P.J., Clohessy, P.A., Grant, G., Pusztai, A. & Golden, B. Faecal calprotectin: non-invasive marker of gastrointestinal inflammation in Salmonella infected rats. *Biochem Soc Trans* 24, 308S (1996).

31. Niewold, T.A., Kerstens, H.H., van der Meulen, J., Smits, M.A. & Hulst, M.M. Development of a porcine small intestinal cDNA micro-array: characterization and functional analysis of the response to enterotoxigenic E. coli. *Vet Immunol Immunopathol* 105, 317-329 (2005).

32. Nikolsky, Y., Ekins, S., Nikolskaya, T. & Bugrim, A. A novel method for generation of signature networks as biomarkers from complex high throughput data. *Toxicol Lett* 158, 20-29 (2005).

33. Ogawa, H*., et al.* Increased expression of HIP/PAP and regenerating gene III in human inflammatory bowel disease and a murine bacterial reconstitution model. *Inflamm Bowel Dis* 9, 162-170 (2003).

34. Ojcius, D.M., Degani, H., Mispelter, J. & Dautry-Varsat, A. Enhancement of ATP levels and glucose metabolism during an infection by Chlamydia. NMR studies of living cells. *J Biol Chem* 273, 7052-7058 (1998).

35. Olsen, S.J*., et al.* The changing epidemiology of Salmonella: trends in serotypes isolated from humans in the United States, 1987-1997. *J Infect Dis* 183, 753-761 (2001).

36. Pedron, T., Thibault, C. & Sansonetti, P.J. The invasive phenotype of Shigella flexneri directs a distinct gene expression pattern in the human intestinal epithelial cell line Caco-2. *J Biol Chem* 278, 33878-33886 (2003).

37. Pellis, L., Franssen-van Hal, N.L., Burema, J. & Keijer, J. The intraclass correlation coefficient applied for evaluation of data correction, labeling methods, and rectal biopsy sampling in DNA microarray experiments. *Physiol Genomics* 16, 99-106 (2003).

38. Rakoff-Nahoum, S., Paglino, J., Eslami-Varzaneh, F., Edberg, S. & Medzhitov, R. Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. *Cell* 118, 229-241 (2004).

39. Reeves, P.G., Nielsen, F.H. & Fahey, G.C., Jr. AIN-93 purified diets for laboratory rodents: final report of the American Institute of Nutrition ad hoc writing committee on the reformulation of the AIN-76A rodent diet. *J Nutr* 123, 1939-1951 (1993).

40. Renton, K.W. Alteration of drug biotransformation and elimination during infection and inflammation. *Pharmacol Ther* 92, 147-163 (2001).

41. Rimoldi, M*., et al.* Intestinal immune homeostasis is regulated by the crosstalk between epithelial cells and dendritic cells. *Nat Immunol* 6, 507-514 (2005).

42. Salzman, N.H*., et al.* Enteric Salmonella infection inhibits Paneth cell antimicrobial peptide expression. *Infect Immun* 71, 1109-1115 (2003).

43. Santos, R.L., Zhang, S., Tsolis, R.M., Baumler, A.J. & Adams, L.G. Morphologic and molecular characterization of Salmonella typhimurium infection in neonatal calves. *Vet Pathol* 39, 200-215 (2002).

44. Santos, R.L*., et al.* Animal models of Salmonella infections: enteritis versus typhoid fever. *Microbes Infect* 3, 1335-1344 (2001).

45. Sass, W., Dreyer, H.P. & Seifert, J. Rapid insorption of small particles in the gut. *Am J Gastroenterol* 85, 255-260 (1990).

46. Schwan, W.R., Huang, X.Z., Hu, L. & Kopecko, D.J. Differential bacterial survival, replication, and apoptosis-inducing ability of Salmonella serovars within human and murine macrophages. *Infect Immun* 68, 1005-1013 (2000).

47. Sorbara, L.R*., et al.* Human immunodeficiency virus type 1 infection of H9 cells induces increased glucose transporter expression. *J Virol* 70, 7275-7279 (1996).

48. Stecher, B*., et al.* Comparison of Salmonella enterica serovar Typhimurium colitis in germfree mice and mice pretreated with streptomycin. *Infect Immun* 73, 3228-3241 (2005).

49. Stokes, N.R., Zhou, X., Meltzer, S.J. & Kaper, J.B. Transcriptional responses of intestinal epithelial cells to infection with Vibrio cholerae. *Infect Immun* 72, 4240-4248 (2004).

50. Ten Bruggencate, S.J., Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L., Katan, M.B. & Van Der Meer, R. Dietary fructo-oligosaccharides and inulin decrease resistance of rats to Salmonella: protective role of calcium. *Gut* 53, 530-535 (2004).

51. Ten Bruggencate, S.J., Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L. & Van der Meer, R. Dietary fructo-oligosaccharides dose-dependently increase translocation of Salmonella in rats. *J Nutr* 133, 2313-2318 (2003).

52. Tian, B., Nowak, D.E., Jamaluddin, M., Wang, S. & Brasier, A.R. Identification of direct genomic targets downstream of the nuclear factor-kappaB transcription factor mediating tumor necrosis factor signaling. *J Biol Chem* 280, 17435-17448 (2005).

53. van Asten, A.J., Koninkx, J.F. & van Dijk, J.E. Salmonella entry: M cells versus absorptive enterocytes. *Vet Microbiol* 108, 149-152 (2005).

54. van Hal, N.L*., et al.* The application of DNA microarrays in gene expression analysis. *J Biotechnol* 78, 271-280 (2000).

55. Vladoianu, I.R., Chang, H.R. & Pechere, J.C. Expression of host resistance to Salmonella typhi and Salmonella typhimurium: bacterial survival within macrophages of murine and human origin. *Microb Pathog* 8, 83-90 (1990).

56. Wallis, T.S. & Galyov, E.E. Molecular basis of Salmonella-induced enteritis. *Mol Microbiol* 36, 997-1005 (2000).

57. Wallis, T.S*., et al.* The nature and role of mucosal damage in relation to Salmonella typhimurium-induced fluid secretion in the rabbit ileum. *J Med Microbiol* 22, 39-49 (1986).

58. Wassell, J., Dolwani, S., Metzner, M., Losty, H. & Hawthorne, A. Faecal calprotectin: a new marker for Crohn's disease? *Ann Clin Biochem* 41, 230-232 (2004).

59. White, P., Liebhaber, S.A. & Cooke, N.E. 129X1/SvJ mouse strain has a novel defect in inflammatory cell recruitment. *J Immunol* 168, 869-874 (2002).

60. Wick, M.J. Living in the danger zone: innate immunity to Salmonella. *Curr Opin Microbiol* 7, 51-57 (2004).

61. Xu, C., Li, C.Y. & Kong, A.N. Induction of phase I, II and III drug metabolism/transport by xenobiotics. *Arch Pharm Res* 28, 249-268 (2005).

62. Zeng, H*., et al.* Flagellin/TLR5 responses in epithelia reveal intertwined activation of inflammatory and apoptotic pathways. *Am J Physiol Gastrointest Liver Physiol* 290, G96-G108 (2006).

63. Zhang, S*., et al.* Secreted effector proteins of Salmonella enterica serotype typhimurium elicit host-specific chemokine profiles in animal models of typhoid fever and enterocolitis. *Infect Immun* 71, 4795-4803 (2003).

Chapter 3

Salmonella induces prominent gene expression in the rat colon

Wendy Rodenburg [1,2,3]

Jaap Keijer [1,2]

Evelien Kramer [1,2]

Susanne Roosing [2]

Carolien Vink [1,3]

Martijn B. Katan [4]

Roelof van der Meer [1,3,5]

Ingeborg M.J. Bovee-Oudenhoven [1,3]

[1] TI Food and Nutrition, Wageningen, The Netherlands

[2] RIKILT Institute of Food Safety, Wageningen, The Netherlands

[3] NIZO Food Research, Ede, The Netherlands

[4] Vrije Universiteit, Institute of Health Sciences, Amsterdam, The Netherlands

[5] Nutrigenomics Consortium, TI Food and Nutrition, Wageningen, The Netherlands

Abstract

Background

Salmonella enteritidis is suggested to translocate in the small intestine. *In vivo* it induces gene expression changes in the ileal mucosa and Peyer's patches. Stimulation of Salmonella translocation by dietary prebiotics fermented in colon suggests involvement of the colon as well. However, effects of Salmonella on colonic gene expression *in vivo* are largely unknown. We aimed to characterize time dependent Salmonella-induced changes of colonic mucosal gene expression in rats using whole genome microarrays. For this, rats were orally infected with *Salmonella enteritidis* to mimic a foodborne infection and colonic gene expression was determined at days 1, 3 and 6 post-infection (n=8 rats per time-point). As fructo-oligosaccharides (FOS) affect colonic physiology, we analyzed colonic mucosal gene expression of FOS-fed versus cellulose-fed rats infected with Salmonella in a separate experiment. Colonic mucosal samples were isolated at day 2 post-infection.

Results

Salmonella affected transport (e.g. Chloride channel calcium activated 6, H^+/K^+ transporting Atp-ase), antimicrobial defense (e.g. Lipopolysaccharide binding protein, Defensin 5 and Phospholipase A2), inflammation (e.g. calprotectin), oxidative stress related genes (e.g. Dual oxidase 2 and Glutathione peroxidase 2) and Proteolysis (e.g. Ubiquitin D and Proteosome subunit beta type 9). Furthermore, Salmonella translocation increased serum IFNγ and many interferon-related genes in colonic mucosa. The gene most strongly induced by Salmonella infection was Pancreatitis Associated Protein (*Pap*), showing >100-fold induction at day 6 after oral infection. Results were confirmed by Q-PCR in individual rats. Stimulation of Salmonella translocation by dietary FOS was accompanied by enhancement of the Salmonella-induced mucosal processes, not by induction of other processes.

Conclusions

We conclude that the colon is a target tissue for Salmonella*,* considering the abundant changes in mucosal gene expression.

Background

Foodborne infections cause a major burden on public health services and represent significant costs in many countries. Salmonella infection is one of the most common and widely distributed foodborne diseases and can be severe in the young, the elderly and patients with weakened immunity. *Salmonella enteritidis* is the most frequently isolated serotype, causing gastroenteritis in most humans and systemic infection in a subpopulation[4,23]. The precise mechanisms of Salmonella-host interaction *in vivo* at early time points after infection are not well known. Insight in pathogen-induced host processes *in vivo* could help to design therapeutic or nutritional strategies for infection prevention. An approach to investigate the effects of a pathogen on host target cells is the use of microarrays that contain the whole genome of the host. This broad approach can reveal biological processes affected by the pathogen. The rat is a good model to study *Salmonella enteritidis*-induced host processes, since salmonellosis in the rat shares many features of human disease [21]. Besides gastroenteritis, a self-limiting systemic infection is observed in rats. The ileum is thought to be the main site of Salmonella invasion in both humans and rats[49]. For this reason we have previously studied Salmonella-induced gene expression in the ileum of rats. This study showed that Salmonella affects only a small number of genes at early time points post-infection [58]. Carbohydrate transport, antimicrobial defense and detoxification were the main affected biological processes. At later time points large numbers of inflammation genes were found to be up-regulated in the ileal mucosa. The colon mucosa is supposed to be protected from Salmonella colonization by the abundant intestinal microflora. Pathogens entering the colon have to compete for nutrients and binding places with the endogenous flora. However, biopsies taken from humans during an infection with nontyphoid Salmonella setorypes suggest that the colon is involved in Salmonella infections[41,46,72]. As most studies focus on the ileum, which is thought to be the most likely site of translocation, only little information is available on Salmonella translocation in the large intestine [24]. Besides indications from studies on biopsies, we have another reason to suspect colonic involvement in Salmonella infection pathology. We have shown earlier that diets supplemented with prebiotics such as fructo-oligosaccharides (FOS), lactulose and inulin consistently increased intestinal Salmonella translocation in rats[3,5,18,69]. As fermentation of FOS, and other prebiotics, occurs in cecum and colon and is very limited in the ileum of humans [2] and rats [22], it is unlikely that prebiotics facilitated translocation of Salmonella at that particular site. This is supported by the absence of ileal inflammation in FOS-fed and Salmonella-infected rats in contrast to profound cecal and colonic inflammation [5].To extend the current limited evidence indicating colonic involvement in Salmonella infection, we used transcriptional profiling to investigate genes and biological processes in the rat colonic mucosa affected by Salmonella. We first studied colonic mucosal gene expression responses at days 1, 3 and 6 after oral Salmonella infection of rats using whole genome microarrays and Q-PCR. In a second infection experiment, we studied whether the increased translocation of Salmonella by dietary FOS was reflected by amplification of Salmonella-induced gene expression changes in the colon.

Methods

Time course infection study
Animals, diet and infection

The experimental protocols were approved by the animal welfare committee of Wageningen University (Wageningen, the Netherlands). Specific pathogen-free male outbred 9 weeks old Wister rats (WU, Harlan, Horst, the Netherlands, n=48 in total), were housed individually in metabolic cages. All animals were kept in a temperature (22-24 C) and humidity (50-60%) controlled room with a 12 h light/dark cycle (lights on from 6 AM to 6 PM). Rats were fed a purified diet during the whole experimental period. The diet contained (per kg) 200 g acid casein, 502 g glucose, 160 g palm oil, 40 g corn oil, 50 g cellulose, 35 g mineral mix (without calcium) and 10 g vitamin mix according to AIN93 recommendations [56]. Diets were low in calcium content (20 mmol $CaHPO_4.2H_2O$/kg) and high in fat content (200 g fat/kg)[6] to mimic the composition of a Western human diet. Food and demineralized drinking water were supplied *ad libitum*. The animals were acclimatized to the housing and dietary condition for 11 days, after which they were orally infected with *S. enteritidis* (clinical isolate, phage type 4 according to international standards; B1214 culture of NIZO food research, Ede, the Netherlands). Salmonella infection was performed by gastric gavage with 1 mL of saline containing $3x10^9$ colony forming units (CFU) of *S. enteritidis*. Non-infected rats received saline only (control). *S. enteritidis* was cultured and stored, as described earlier [68]. Fresh fecal samples were collected on days 1, 2, 3 and 6 post infection (p.i.) and analyzed for viable Salmonella by plating 10-fold dilutions in sterile saline on Modified Brilliant Green Agar (Oxoid, Basingstoke, UK) and incubating aerobically overnight at 37°C. Sulphamandelate (Oxoid) was added to the agar plates to suppress swarming bacteria, such as Proteus species. The detection limit of this method was 10^2 CFU/g fecal wet weight. Total 24 h urine samples were collected from the day before oral infection of the rats until day 6 after infection. Urines were preserved by adding oxytetracycline to the urine collection vessels of the metabolic cages, and analyzed for the nitric oxide metabolites nitrite and nitrate (summed as NO_x) by a colorimetric method (Nr. 1746081; Roche diagnostics, Mannheim, Germany).

Rats were sacrificed on day 1, 3 or 6 post infection and control (n=8 rats per treatment and per time point). Rats were killed by carbon dioxide inhalation. Blood was collected by orbita puncture. Blood was coagulated for 30 minutes at room temperature, cooled to 4°C and centrifuged 20 minutes by 3000 g. Serum was collected and frozen at -80°C. The mesenteric lymph nodes (MLN), spleen and liver were excised aseptically, weighed, homogenized (Ultraturrax Pro200, Pro Scientific Inc. Oxford, CT) in sterile saline, serially diluted, and plated to culture for Salmonella, as described above. The detection limit was 10^2 CFU/g tissue. To obtain colonic mucosa, the colon was taken out, longitudinally opened and colonic contents removed by a quick rinse in 154 mM KCl. The mucosa was scraped off using a spatula. The scrapings were immediately frozen in liquid nitrogen and stored at -80°C for RNA extraction.

RNA isolation

Colon scrapings were homogenized in liquid N_2 using a mortar and pestle cooled with liquid N_2 (Fisher Emergo, Landsmeer, The Netherlands). Total RNA was isolated from these homogenates using TRIzol reagent (Invitrogen, San Diego, CA) according to the manufacturer's instructions. Total RNA was purified using Rneasy columns (Qiagen, Chatsworth, CA). Absence of RNA degradation was checked on a 1% TBE/agarose gel after 1 hour incubation at 37°C. The purity and concentration were measured with the Nanodrop (Isogen Life Science, Maarssen, The Netherlands). OD A_{260}/A_{280} ratios were all between 2.08 and 2.10 indicating RNA of high purity.

Analysis of mRNA expression by Oligo Arrays

For microarray hybridization, equal amounts of RNA of each animal were pooled per treatment group. Arrays were performed in duplicate. For this, RNA pools were split and separately reverse transcribed and labeled with Cy5. A standard reference sample, consisting of a pool of all colonic RNA was labeled with Cy3. For each oligo array, 35 µg of total RNA was used to make Cy5 or Cy3 labeled cDNA. Total RNA was mixed with 4 µg T21 primer, heated at 65°C for 3 min (RNA denaturation) followed by 25 C for 10 min (primer annealing). cDNA was synthesized by adding 5x first strand buffer (Invitrogen), 10 mM DTT, 0.5mM dATP, 0.5 mM dGTP, 0.5 mM dTTP, 0.04 mM dCTP, 0.04 mM Cy5-dCTP or Cy3-dCTP, 1.2U RnaseOUT and 6U SuperScript II Reverse Transcriptase to a total volume of 62.5 µL. The reaction was incubated at 42°C for 2 h. Purification, precipitation and denaturation of the labeled cDNA were performed as described earlier [70].

The 44K rat whole genome Agilent array (G4131A, Agilent Technologies, Inc. Santa Clara, CA) used consists of 44290 60-mer rat oligonucleotides, including ~3000 control spots. The Cy5 labeled cDNAs of the Salmonella infected groups and the non-infected groups were mixed 1:1 with the Cy3 reference labeled cDNA, mixed with 2x hybridization buffer (Agilent Technologies) and 10x control targets (Agilent Technologies) and hybridized for 17 hours at 60°C in Agilent hybridization chambers in an Agilent hybridization oven rotating at 4 rpm (Agilent Technologies). After hybridization the arrays were washed with an SSPE wash procedure (Agilent Technologies) and scanned with an Agilent Microarray Scanner (Agilent Technologies).

Data analysis

Signal intensities for each spot were quantified using Feature Extraction 8.1 (Agilent Technologies). The data of the time course infections study are available in Additional File 5 at the BMC website (http://www.biomedcentral.com) and have been deposited in NCBIs Gene Expression Omnibus [50] and are accessible through GEO Series accession number GSE7496. Median density values and background values of each spot were extracted for both the experimental samples (Cy5) and the reference samples (Cy3). Quality check was performed for each microarray using both LimmaGUI package in R from Bioconductor [73] and Microsoft Excel. Data was exported into GeneMaths XT (Applied Maths, Sint-Martens-Latem, Belgium) for analysis. We discarded spots with an average intensity, over all arrays, of Cy5 lower than 2-fold above average background. Then, the Cy5 intensities were normalized against the Cy3 reference as described before [54]. The gene expressions of duplicate arrays were averaged. Array data of non-

infected rats, killed on section day 1 and 6 were highly comparable and could therefore be considered as one group and were averaged. For unknown reason, arrays of non-infected rats killed on day 3 showed reduced expression of 14 mast cell protease genes when compared with non-infected rats of both days 1 and 6, which were highly comparable. Therefore, we decided not to include the non-infected rats of day 3. Cluster analysis and Principle component analysis were performed using GeneMaths XT. Infected/control ratio's between 0-1 were expressed as the negative inverse (-1/value) for easier interpretation. Genes that changed more than 2-fold in comparison with controls at one of the time points studied were selected for pathway analysis. Pathway analysis was performed using two pathway programs, MetaCore (GeneGo Inc, St. Joseph, MI)[16] and ErmineJ [40], using Agilent gene annotation (Agilent Technologies, version 20060331). Processes were identified using statistical over-representation in both pathway programs. Since only 40% of the genes were annotated to GO processes in both pathway programs, processes with a p-value<0.001 were manually supplemented with non-annotated genes with FC>2 using biological databases (BIOcarta, SOURCE, GenMAPP, KEGG) and scientific literature.

Analysis of mRNA expression by Real-time Quantitative RT-PCR

Real-time Quantitative RT-PCR (Q-PCR) was performed on individual samples (n=8 per group). 1 μg of RNA of all individual samples was used for the cDNA synthesis using the iScript cDNA synthesis kit of Bio-Rad Laboratories (Veenendaal, The Netherlands). Real-time reactions were performed by means of the iQ SYBR Green Supermix of Bio-Rad using the MyIQ single-color real-time PCR detection system (Bio-Rad). Each reaction (25 μl) contained 12.5 μl iQ SYBR green supermix, 1 μl forward primer (10 μM), 1 μl reverse primer (10 μM), 8.5 μl RNase-free water and 2 μl diluted cDNA. The following cycles were performed 1x 3 min at 95°C, 40 amplification cycles (40x 10 s 95°C, 45 s 60°C), 1x 1 min 95 °C, 1x 1 min 62 °C and a melting curve (80x 10 s 55 °C with an increase of 0.5 °C per 10 s). A negative control without cDNA template was run with every assay. The optimal melting point of dsDNA (Tm) and the efficiency of the reaction were optimized beforehand. Data were normalized against the reference genes Ribosomal protein S29 (*Rps29*), ADP-Ribosylation Factor 1 (*Arf1*) and β-actin. *Rps29* and *Arf1* were chosen on the basis of microarray data which showed similar expression levels for all microarrays, β-actin was chosen as this is a well accepted reference gene. Primers were designed using Beacon designer 4 (Premier Biosoft International, Palo Alto, CA). For sequences see supplemental table 3.1. A standard curve for all genes including reference genes was generated using serial dilutions of a pooled sample (cDNA from all reactions). mRNA levels were determined from the appropriate standard curve. Samples with mRNA levels below the lowest standard value, and thus below detection level, were given half the value of this lowest standard. Analysis of all individual samples was performed in duplicate.

Serum Interferon Gamma

The serum Interferon Gamma (IFNγ) concentration of individual rats was determined by an enzyme-linked immunosorbent assay (ELISA) specific for rats (Biosource International, Camarillo, CA) according to the manufacturer's protocol.

Dietary infection study

Animals, diet and infection

A dietary intervention was performed to study the effect of FOS on *S. enteritidis*-induced gene expression. Specific pathogen-free male outbred Wister rats (8 weeks old, mean body weight of 253 g; n=48 in total) were housed as described above (time course infection study). Rats were fed the same diet as described above. The experimental diets both contained 20 g/kg cellulose and were supplemented with either 60 g/kg FOS (purity 93%; Raftilose P95, Orafti, Tienen, Belgium) or additional 60 g/kg cellulose as described earlier [68]. Animals were fed restricted quantities (14 g/day) of the purified diet. Restricted food intake was necessary to prevent differences in food consumption and hence differences in vitamin and mineral intake as observed earlier in FOS interventions [5]. After an adaptation period of 14 days, rats were orally infected with 4×10^8 CFU of *S. enteritidis* or control-treated as described above. On day 2 p.i., 12 infected FOS-fed rats, 12 infected rats fed the cellulose diet, and 12 control-treated non-infected rats fed the cellulose diet were sacrificed to obtain colonic mucosal RNA. Two additional groups of rats fed either FOS (n=6) or the cellulose diet (n=6) and infected with Salmonella were kept until day 8 p.i. for determination of urinary NO_x excretion in time as described above.

Analysis of mRNA expression by Oligo Arrays and Real-time Quantitative RT-PCR

RNA isolation and analysis of mRNA expression by microarray (pooled samples) and Q-PCR (n=12 per treatment group) were performed as described above. Arrays were scanned with a Scanarray Express HT scanner (Perkin Elmer). Signal intensities for each spot were quantified using ArrayVision 8.0 (GE Healthcare life sciences). Data analysis was performed as described above. The data of the dietary infection study available in Additional File 6 at the BMC website (http://www.biomedcentral.com) and have been deposited in NCBIs Gene Expression Omnibus [50] and are accessible through GEO Series accession number GSE7472.

Statistical analysis

Results are expressed as median or mean depending on normality of distribution as indicated. We used Prism 4 for all statistics (Prism 4, GraphPad software Inc., San Diego, CA). Data was analyzed using the Student's t-test (two-sided). Non-normally distributed data was analyzed using the non-parametric Mann-Whitney U test (two sided). Differences were considered statistically significant when $p < 0.05$.

Results

Time course infection study
General infection characteristics
In agreement with previous studies, food consumption and growth of the Wistar rats were not affected by Salmonella infection[6]. Salmonella translocation to mesenteric lymph nodes was observed at days 1, 3 and 6 (table 1). This implies that at day 1, Salmonella has already crossed the intestinal barrier. In agreement with previous studies[6,53], Salmonella was detected in the spleen at days 3 and 6 and in the liver at day 6 (table 1). Urinary NO_x excretion, a parameter of systemic infection, was found to be increased from day 3 onwards (figure 1a).

Table 1. Viable Salmonella counts in feces, mesenteric lymph nodes, spleen, and liver of rats 1, 3 and 6 days post infection.

		Salmonella (logCFU/g)[a]		
	Control	Day 1 p.i.	Day 3 p.i.	Day 6 p.i.
Feces	N.D.	7.22±0.19	5.92±0.24	6.04±0.32
MLN	N.D.	3.38±0.43	5.85±0.13	5.44±0.05
Spleen	N.D.	N.D	3.20±0.33	3.49±0.05
Liver	N.D.	N.D	N.D	2.45±0.12

[a] The rats were orally infected with 3×10^9 colony-forming units of *S. enteritidis* or control treated. Salmonella counts are expressed in log values as means ± SEM (n=8). N.D. = not detected.
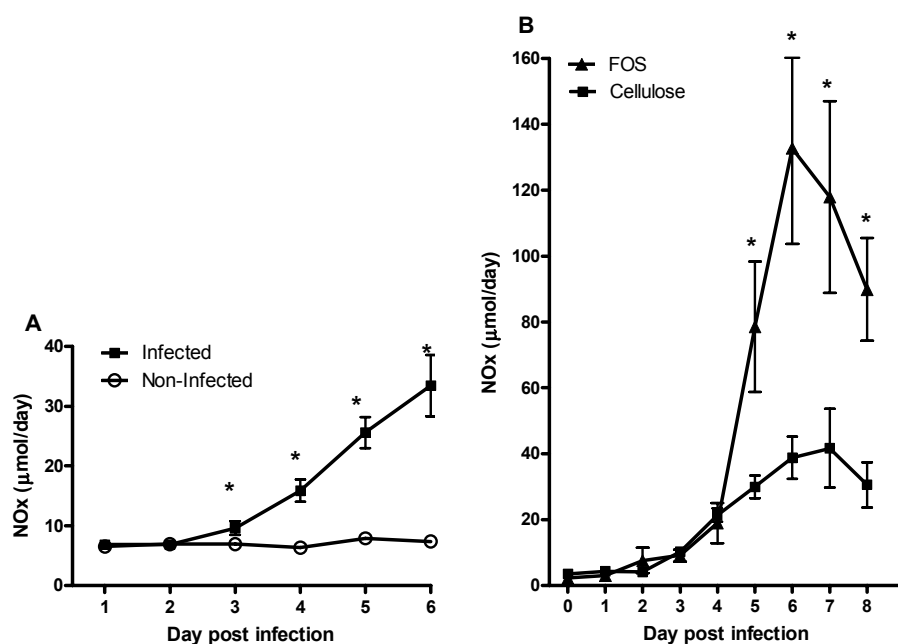
Figure 1. Sum of urinary nitrate and nitrite (NO_x) excretion in the non-infected (o), infected (■) groups of the time course infection study (A). And the urinary NO_x excretion in the cellulose infected (■) and in the fructo-oligosaccharide (FOS) infected (▲) groups in the dietary infection study (B). Infected rats were orally challenged with S. *enteritidis* on day 0. Results are expressed as mean ± SEM (n=8 in the time course infection study and n=6 in the dietary infection study). * $p<0.05$.

Salmonella-induced processes in colon mucosa
To identify Salmonella-regulated processes, microarray-based gene expression profiling of colonic mucosa at days 1, 3 and 6 days p.i. was performed. The arrays contained 44000 spots of which 32783 spots exceeded >2 times the background value and were included in the analysis.

Salmonella changed the expression of 330 genes >2-fold at least at one of the three time points studied. At days 1 and 3 p.i. comparable numbers of genes (70 and 57 genes, respectively) were affected by Salmonella infection in comparison with non-infected rats, while at day 6 approximately four times more genes were affected (figure 2). This corresponded with progression of the infection as observed by the organ cultures and urinary NO_x excretion as mentioned above. At all time points studied, most genes showed increased expression upon Salmonella infection, whereas only a small percentage of total regulated genes were down-regulated (10% at day 1, 27% at day 3, 5% at day 6 figure 2). The genes that changed more than 2-fold at any time point (FC> 2 infected/ non-infected) were classified into biological processes according to gene ontology terminology [16,40]. Not all genes are annotated to GO processes. Forty percent of the genes on the array were annotated to GO processes. Therefore we manually supplemented the significant processes (p<0.001) with the remaining significant genes using biological databases and scientific literature. To prevent the occurrence of false positive genes, and over-interpretation of biological processes affected by Salmonella, we focused on biological processes with at least three genes exceeding the cut-off FC>2.0. Additionally, we observed that the genes within one biological process showed comparable patterns of expression (table 2), which strongly indicates that these processes are truly affected by Salmonella.

Figure 2. The number of differentially expressed genes with a fold change greater than 2 in colon mucosa of rats at days 1, 3 or 6 after oral infection with Salmonella or control treatment.

We focused on the early Salmonella-induced gene expression changes occurring at days 1 and 3 p.i. Presumably, these early modulated genes are more related to Salmonella-induced primary changes than gene expression at day 6 which is a secondary result of Salmonella-induced inflammation. Genes affected >2-fold on day 1 and/or day 3 p.i. that could be related to a biological process are shown in table 2. The biological processes that contained 3 or more modulated genes were transport, oxidative stress, immune response, antimicrobial defense, inflammatory response, interferon pathways and proteolysis. For more insight into these processes, genes that changed >2-fold at day 6 p.i. and also showing a >1.5-fold induction at day 1 or 3 p.i. were also added to this table. Genes that changed >2-fold on day 6 p.i. only are available in Additional File 1 at the BMC website (http://www.biomedcentral.com). The gene most affected by Salmonella infection in the colon was pancreatitis associated protein (*Pap*), showing 11, 45 and 114 fold induction at days 1, 3 and 6 respectively.

Table 2. Processes regulated in colon by Salmonella at days 1, 3 and 6 after oral Salmonella infection.

Gene Name		Gene symbol	Sequence ID	Fold Change infected vs non-infected rats on different days p.i.			
				Time course infection study[a]			Dietary infection study[b]
				Day 1	Day 3	Day 6	Day 2
Transport							
Chloride channel calcium activated 6		*Clca6*	NM_201419	**2,3**	**2,2**	**3,7**	*2,3*
Calcium channel, voltage-dependent, alpha 1I subunit		*Cacna1i*	NM_020084	**2,2**	1,5	-	*1,2*
Solute carrier family 4, member 1 (Slc4a1) , anion exchanger		*Slc4a1*	NM_012651	**2,0**	1,7	-	*1,0*
Atp-ase, H+/K+ transporting, nongastric, alpha polypeptide		*Atp12a*	NM_133517	**2,8**	-	**2,0**	*1,1*
Solute carrier family 20 (phosphate transporter), member 1		*Slc20a1*	NM_031148	**2,1**	-	-	*-1,2*
Solute carrier family 15 (oligopeptide transporter), member 1		*Slc15a1*	NM_057121	**2,0**	**1,5**	**1,6**	*0,9*
Transporter 1, ATP-binding cassette, sub-family B		*Mdr/ Tap1*	NM_032055	-	1,8	**2,8**	*1,4*
Oxidative stress							
Dual oxidase 2		*Duox2*	NM_024141	1,9	**2,4**	**2,8**	*1,6*
Glutathione peroxidase 2		*Gpx2*	NM_183403	-	**2,3**	**3,0**	*2,2*
Xanthine dehydrogenase		*Xdh*	NM_017154	-	1,8	**2,5**	*1,4*
Immune response							
Rat class III Fc-gamma receptor		*Fcgr3*	M64368	**2,1**	1,6	-	*ND*
Immunoglobulin superfamily, member 4		*Igsf4d*	XM_340958	**2,1**	1,6	1,4	*1,1*
Rat MHC class I truncated cell surface antigen		*RT1-Aw2*	M10094	**2,0**	-	1,9	*1,0*
Interleukin enhancer-binding factor 1		*Ilf1*	XM_221212	**2,0**	1,5	-	*1,1*
Colony stimulating factor 2 (granulocyte-macrophage)		*Csf2*	XM_340799	**2,0**	1,6	-	*1,2*
Interleukin 1 alpha		*Il1a*	NM_017019	1,8	**2,0**	**2,3**	*1,4*
Interleukin 1 beta		*Il1b*	NM_031512	-	**2,1**	**4,0**	*2,6*
TRAF2 binding protein		*T2bp*	NM_001014044	1,8	**2,8**	**4,4**	*2,4*
Toll-like receptor 2		*Tlr2*	NM_198769	-	1,5	**2,3**	*1,3*
Antimicrobial defense							
Lipopolysaccharide binding protein		*Lbp*	NM_017208	1,9	1,8	**2,2**	*1,3*
Defensin 5 precursor (Enteric defensin)		*Rd5*	XM_214386	-1,9	-1,6	-1,6	*ND*
Phospholipase A2, group IIA (platelets, synovial fluid)		*Pla2g2a*	NM_031598	**3,4**	**5,2**	**10,5**	*7,3*
Inflammatory response							
Pancreatitis-associated protein		*Pap*	NM_053289	**11,4**	**44,6**	**114,2**	*17,7*
Tissue-type transglutaminase		*Tgm2*	NM_019386	-	**2,3**	**4,9**	*1,8*
Regenerating islet-derived 3 gamma		*Reg3g*	NM_173097	-	**2,3**	**4,3**	*1,9*
Nitric oxide synthase 2, inducible		*Nos2*	NM_012611	-	1,6	**4,0**	*ND*
S100 calcium binding protein A8	Calprotectin	*S100a8*	NM_053822	-	1,8	**2,4**	*1,4*
S100 calcium binding protein A9		*S100a9*	NM_053587	-	1,7	1,9	*1,3*
Interferon							
Interferon-induced guanylate-binding protein 1		*Gbp1*	XM_221883	**2,4**	1,9	**2,2**	*1,3*
Interferon gamma inducible protein		*Ifi47*	NM_172019	1,7	**2,7**	**7,3**	*2,6*
Guanylate binding protein 2, interferon-inducible		*Gbp2*	NM_133624	1,6	**2,4**	**3,1**	*1,7*
Interferon-induced protein		*Ifit2*	NM_001024753	1,5	1,7	**3,4**	*1,3*
Interferon-stimulated protein		*G1P2*	XM_216605	-	1,6	**4,1**	*1,5*
Immunity-related GTPase family, M		*Irgm*	NM_001012007	-	1,8	**3,7**	*1,4*
Signal transducer and activator of transcription 1		*Stat1*	NM_032612	-	1,7	**3,6**	*1,8*
Interferon regulatory factor 7		*Irf7*	XM_215121	-	1,5	**2,6**	*ND*
Alpha-interferon		*Ifna*	XM_233145	-	**2,0**	-	*ND*
Proteolysis							
Ubiquitin D		*Ubd*	NM_053299	1,7	**2,5**	**15,2**	*3,4*
Proteosome (prosome, macropain) subunit, beta type 9		*Psmb9*	NM_012708	-	**2,0**	**3,7**	*2,0*

Protease, serine, 22	*Prss22*	XM_220222	-	**2,0**	**2,6**	*1,6*
Potential ubiquitin ligase	*Herc6*	XM_342700	-	1,7	**3,3**	*1,5*
Proteasome (prosome, macropain) subunit, beta type 10	*Psmb10*	XM_214687	-	1,5	**2,1**	*1,5*

[a] Values in bold exceed cut-off value FC>2 or FC<-2. Values -1,5 <FC< 1,5 are indicated by (-).

[b] Fold Change infected vs non-infected rats fed a cellulose diet at day 2 p.i (obtained from the dietary infection study). All fold changes are shown. Genes not detected in this independent study are indicated by ND.

Seventy genes changed>2-fold at day 1 p.i., of these genes 7 encoded for transporters and 5 genes encoded for immune response proteins (table 2). At day 3 p.i., 57 genes showed FC>2, again including genes encoding for immune response proteins. Induced expression of Interleukin 1β and 1α indicates activation of an inflammatory response. Induction of dual oxidase 2 and glutathione peroxidase 2 suggest oxidative stress in the colonic mucosa. At day 6 more than 200 genes were induced more than 2-fold in infected mucosa compared with non-infected mucosa (Additional File 1 at BMC website). Most of these genes were related to immune and inflammatory responses. Processes related to inflammation-induced damage and repair, such as connective tissue remodeling and chemo-attraction also showed clear induction at day 6 p.i. To exclude the possibility that the observed changes were due to cellular changes of the mucosa, we analyzed expression differences of cell-type specific genes [26,37] (supplemental table 3.2). As transporters are most likely expressed by enterocytes [1], we examined expression of enterocyte specific genes (*Fabp2, Vil2, Alpi2*). These genes showed diverse regulation, indicating that the increased expression of transporters at day 1 is not due to altered enterocyte composition in the mucosal samples. A similar observation was found for Goblet cell specific genes (*Muc2, Muc3, Tff1, Tff3*) and Paneth cell specific genes (*Rd5* and *Pla2g2a*). Expression of leukocyte specific genes was not altered at early timepoint, a mild increase was observed at day 6 p.i. Together this indicated that the observed Salmonella induced gene expression changes did not result from changes in cellular composition of the mucosa. This is in agreement with histology results from earlier Salmonella infection experiments, showing no or only minor deviations in intestinal mucosal architecture from healthy control slides (data not shown). This is further supported by the relatively constant expression of a group of well known housekeeping genes (supplemental table 3.2). The largest group of related genes induced by Salmonella infection in colon mucosa is related to interferon pathways as more than 20 IFNγ-regulated genes showed increased expression at at least one time point studied. The IFNγ-induced gene expression was most prominent at day 6 p.i, but already from day 1 onwards induction of several IFNγ-inducible GTPases (*Gbp1, Gbp2, Ifi47, Ifit2*) was seen (table 2). Furthermore at day 3 p.i. (table 2) induction of two members of the IFNγ- signaling pathway (*Stat1* and *Irf7*) was observed. Despite induction of many interferon-related genes, increased expression of IFNγ mRNA itself could not be detected (changed 1.1-fold at days 1 and 3, and 1.3-fold at day 6 p.i.). IFNγ protein concentrations were measured in individual serum samples. IFNγ was not detected in serum of non-infected rats (all timepoints) and at the first day after Salmonella administration to rats. However, from day 3 p.i. serum IFNγ increased (figure 3). The serum IFNγ most probably originated from peripheral immune activation, as the increase in serum IFNγ followed the same trend as the increase in Salmonella CFU's in peripheral organs (table 1). In the time course infection study, the kinetics of urinary NO_x excretion are reflected by *Nos2* gene expression in

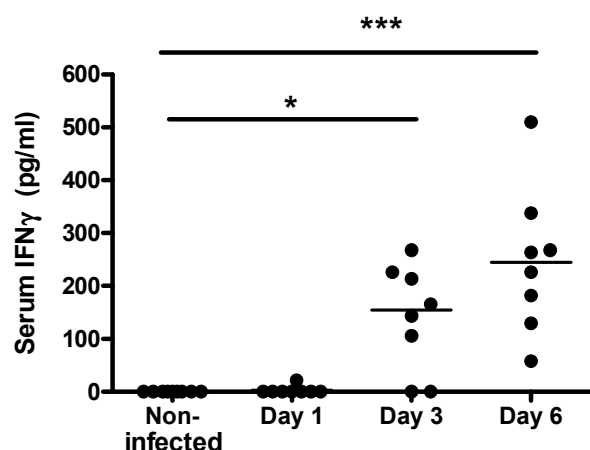colonic mucosal with a small 1.6-fold induction at day 3 p.i. and a 4 fold induction at day 6 p.i. (table 2).

Figure 3. Serum IFNγ levels before and after infection (days 1, 3 and 6 p.i.). Each dot represents an individual rat. Group medians are presented by a black line. * $p < 0.05$, ***$p < 0.001$.

Q-PCR confirmation of Salmonella-induced gene-expression

To determine inter-individual variation in gene expression within treatment groups, RNA from the colon of individual animals was analyzed by Q-PCR. We chose individual confirmation of *Stat1* and *Ifi47* to gain insight in inter individual interferon response as we also focused on the individual protein levels of IFNγ. Confirmation of *Pap* was chosen to obtain insight in the individual kinetics of the most strongly induced gene in colon mucosa at all time points. Q-PCR analysis showed rather large inter-individual variation among the outbred rats. *Pap* expression levels in the non-infected colonic mucosa were near detection level, which made it difficult to determine precise fold changes. Nevertheless, the Q-PCR analysis of the three genes examined clearly confirmed the gene expression changes observed in the microarray analysis (figure 4). To further validate the array data of the time course infection study we compared the gene expression changes of day 1 and 3 p.i. with gene expression data obtained from the independent dietary infection study at day 2 p.i. (table 2). Similar biological processes were induced at early timepoints in both studies. At individual gene expression level several transporter genes (*Cacna1i, Slc4a1, Slc15a1*) and immune response genes (*Igsf4d, Ilf1, Csf2*) showed no overlap possibly due to infection kinetics. However gene expression results of day 3 and day 2 p.i. largely overlapped (table 2).
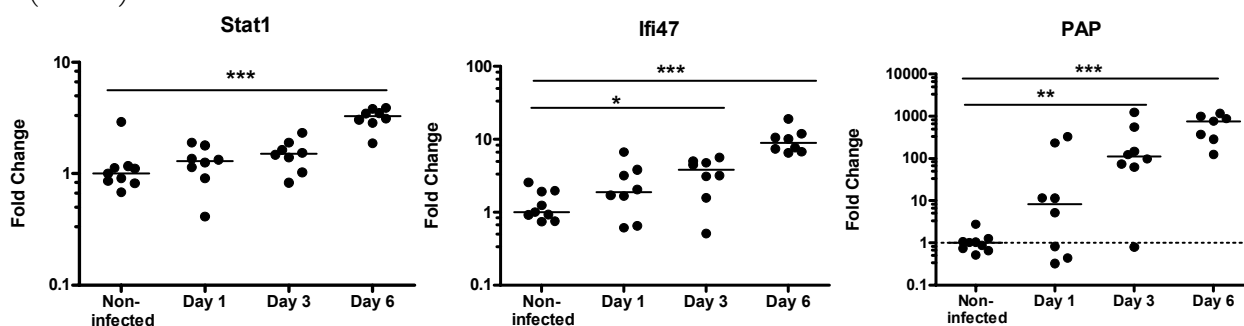

Figure 4. Individual expression of two interferon-related genes and *Pap* in the colon mucosa at different time points after Salmonella infection or control treatment. Genes expression is quantified by Q-PCR, using *Rps26* as reference gene (using *Arf1* as reference gene showed similar results; data not shown). Each dot represents an individual rat. Dotted line indicate lowest mRNA standard. Medians are presented by a black line. The median value of the uninfected group was set at 1.0. Y-axis is at $\log_{10}$ scale. * $p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Dietary infection study

General infection characteristics

Food consumption and growth of the Wistar rats on both cellulose and FOS diet were similar, before and after infection. The section was performed on day 2 p.i. because similar effects on gene expression at days 1 and 3 were observed in the time course infection study. Furthermore we were interested in the primary responses as we expect that diet will mainly influence early events such as attachment to the mucosa and translocation of the pathogen. These events occur for the most part luminal where direct interaction between dietary components, pathogens and mucosa take place, while later phases merely reflect secondary infection and inflammation responses. At day 3 p.i. the serum IFNγ concentration and the amount of translocated Salmonella in the spleen are already high, which indicates systemic infection. At day 1 p.i. no serum IFNγ or Salmonella translocation to the spleen was observed. We chose day 2 p.i. as this seems the appropriate time point to study the effects of FOS on early Salmonella-induced changes. Salmonella colonization was quantified by determination of colony-forming units (CFU/g) in fresh fecal samples with time. At day 1 Salmonella levels were not significantly different between cellulose and FOS-fed animals (7.23 ± 0.11 and 7.10 ± 0.22, respectively). At day 2 FOS-fed rats had more Salmonella in feces than their cellulose counterparts (7.25 ± 0.25 and 6.53 ± 0.25, respectively; $p<0.05$). Salmonella translocation to mesenteric lymph nodes and spleen was not significantly different in the FOS group compared to the cellulose group when quantified by CFU. Viable Salmonella counts in MLN were 5.96 ± 0.08 in the cellulose group and $6.19 \pm 0.10 \log_{10}$ CFU/g in the FOS group. Numbers in spleen were 2.85 ± 0.14 (cellulose) and $2.98 \pm 0.16 \log_{10}$ CFU/g (FOS). Counts in liver were under the detection limit of 10^2 CFU/g tissue in the cellulose and FOS group. These numbers are comparable to those observed in the time course infection study at day 3 and highly similar to numbers observed in earlier studies which showed increased translocation in FOS-fed rats at later time points after infection [68]. To observe long term effects of FOS on Salmonella translocation in this study, urinary NO_x excretion with time was determined in additional groups of rats. Urinary NO_x excretion of FOS-fed rats increased to 132 µmol/d at day 6 p.i. and started to decline towards baseline levels thereafter (figure 1B). Peak urinary NO_x excretion of infected rats fed the cellulose diet was just one third of the level reached by the infected rats fed the FOS diet, i.e. 41 µmol/day (figure 1B). The NO_x values for the cellulose diet are similar to those obtained in the time course infection study (figure 1A). The kinetics of urinary NO_x excretion were similar in both diet groups, but total infection-induced urinary NO_x excretion was higher in the FOS group indicating enhanced Salmonella translocation.

Effect of dietary FOS on Salmonella-induced mucosal genes in colon mucosa

The rats of the time course infection study and the rats in the dietary infection study on cellulose diet showed a comparable urinary NO_x excretion and thus Salmonella translocation response. Despite the fact that the two studies were separately performed and different time points were studied, the identified biological processes affected by Salmonella at day two p.i. were comparable to processes observed at days 1 and 3 p.i. Furthermore both studies showed that more genes were up-regulated than down-regulated by Salmonella. For detailed analysis, we focused on the most robust genes, i.e. genes that showed similar Salmonella induced regulation in the two

independent studies. We choose a threshold of FC<1.5 for both studies, which is less stringent than the threshold we choose for analysis within one study (FC<2.0). We feel that this is legitimate, as genes with small but similar regulation in two completely independent studies are less likely to be selected by chance. Thirty-one genes fulfilled this criterion, 26 were up-regulated and 5 down-regulated. Eighteen of the up-regulated genes were categorized to the same processes found to be modulated by Salmonella in the colonic mucosa in the time course infection study, i.e. the transporter *Clca6*, the oxidative stress genes *Gpx2* and *Duox2*, the immune response genes *Il1b and T2bp*, the antimicrobial defense gene *Pla2g2a* the inflammatory response genes *Pap, Tgm2 and Reg3g*, the interferon related genes *Ifi47, Gbp2, Iigp2, P47Iigp, Stat1, G1p2* and the proteasome related genes *Psmb9, Prss22, Psmb10, Ubd (*Table 2). The other 8 up-regulated genes which could not be grouped to a specific process were Palmitoyl-protein thioesterase, Schlafen 3 (*Slfn3*), Tripartite motif protein 15 (*Trim15*), Aquaporin 3 (*Aqp3*) and four unknown genes. The 5 down-regulated genes were Heat shock protein 70kD 1A (*Hspa1a*), Resistin like alpha (*Retnla*), Resistin like gamma *(Retnlg)*, Collectin sub-family member 10 (*Colec10*) and Mammalian suppressor of Sec4 (*Mss4*). Not all processes that were identified in the time course infection study at both days 1 and 3 p.i. were confirmed in the dietary infection study at day 2 p.i. (table 2). This was the case for two processes, namely transport (*Cacna1i, Slc4a1, Atp12a, Slc15a1*) and immune response (*Igsf4d, RT1, Ilf1, Csf2, Il1a*). Furthermore the antimicrobial defense gene *Lbp* and two interferon pathway genes (*Gbp1, Ifit2*) were not confirmed.

To examine whether our choice for FC>1.5 was legitimate, we studied whether application of threshold FC>1.3 and FC>1.7 resulted in identification of the same processes as identified with FC>1.5. The general picture of processes affected was the same for FC>1.5 and FC>1.7. However, with FC>1.3 more genes could be included in processes identified with FC>1.5, such as the interferon response and proteolysis (data not shown). However, many other genes could not be grouped into (new) specific biological processes, indicating that a cut-off FC>1.3 might be too flexible and results in introduction of false positive processes, probably not related to the treatment. Therefore, we choose FC>1.5 for further analysis. To investigate the effects of FOS on Salmonella infection in the colon, we studied the expression of Salmonella-induced colonic mucosal genes in infected rats fed the cellulose diet versus infected rats fed the FOS supplemented diet. The five genes that were consistently downregulated by Salmonella in both studies (*Hspa1a, Retnla, Retnlg, Colec 10 and Mss4*) were not further influenced by FOS (equal gene expression in cellulose- and FOS-fed infected rats). For initiating early mucosal events after Salmonella infection (e.g. chemo attraction of inflammatory cells) increases in epithelial gene expression may be more important than decreases[15,34]. We focused on the 26 genes which showed a consistent increase in gene expression after Salmonella infection of FC>1.5 in both studies. All 26 genes consistently induced by Salmonella infection in the colon mucosa showed a further up-regulation in colon mucosa of Salmonella infected rats fed FOS (figure 5). The effect of FOS on the cluster of Salmonella affected genes was statistically significant. To asses the inter-individual gene expression in the dietary infection study we selected genes from several Salmonella modulated process for individual Q-PCR confirmation: *Clca6, Gpx2, Il1b, Pla2g2a, Pap, Tgm2, Stat1, Gbp2* and *Ifi47*. Q-PCR of the selected genes in individual samples showed high inter-
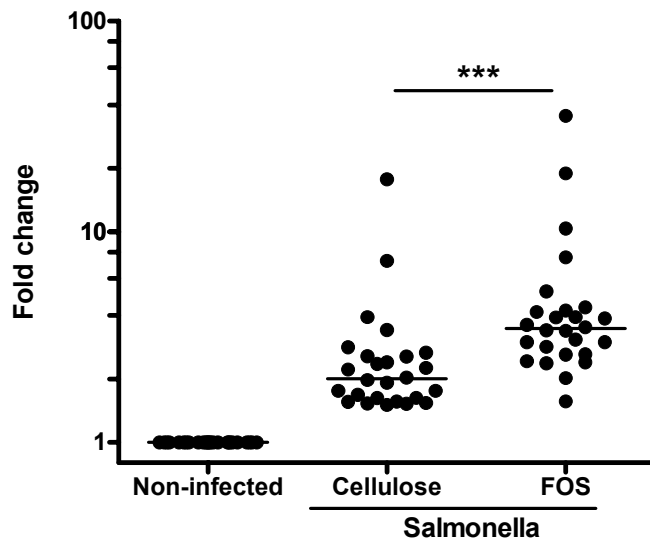
Figure 5. Expression level of the most consistent Salmonella-target genes in the colon mucosa of rats fed a cellulose diet or a FOS diet. The gene expression is obtained from micro array analysis of pooled colonic mucosa samples collected at day 2 post-infection. Each dot represents a gene. The median value of each gene in the uninfected group is set to 1.0. Y-axis is at $\log_2$ scale. ***$p < 0.001$.

individual variation but confirmed the fold changes of the microarray study using pooled samples (table 3). The confirmed Salmonella induced gene expression changes were significant ($p < 0.05$) for 7 of the 9 genes, except for *Tgm2* (p =0.09) and *Stat1* (p=0.08). Examination of FOS-fed versus cellulose-fed infected groups on individual gene level showed a significant increase of *Clca6* and *Pla2g2a*. Expression of *Gpx2, Il1β* and *Tgm2* was >1.5-fold increased by FOS feeding in comparison to cellulose feeding but this was not statistically significant. The genes *Pap, Stat1, Ifi47* and *Gbp2* showed non-significant and small increases of 1.1-1.4 fold.

Table 3. Q-PCR analysis of Salmonella-induced colonic mucosal gene expression of rats on a cellulose or a FOS diet (day 2 p.i.).

	Relative gene expression [a]		
	Non-infected [b]	Infected	
Gene symbol		Cellulose	FOS
Clca6	1 (0.9-1.3)	1.6 (1.4-2.2)	3.7 (2.8-4.1)
Pla2g2a	1 (0.8-1.3)	4.6 (2.3-7.1)	8.8 (6.1-21.1)
Gpx2	1 (0.8-1.2)	2.1 (1.7-2.4)	3.2 (1.3-4.8)
Il1b	1 (0.6-1.1)	2.0 (1.3-3.4)	3.8 (0.6-5.4)
Tgm2	1 (0.5-1.2)	1.6 (0.7-4.4)	10.2 (0.8-20.6)
Pap	1 (0.4-2.4)	236 (68-326)	288 (13-1162)
Stat1	1 (0.7-1.2)	1.2 (1.1-2.2)	1.5 (0.7-4.0)
Ifi47	1 (0.8-1.2)	3.9 (2.1-6.4)	5.6 (1.6-18.4)
Gbp2	1 (0.8-1.4)	2.0 (1.1-3.2)	2.1 (0.5-9.7)

[a] The expression of genes is analyzed by Q-PCR, using *Rps26* as reference gene (using *Arf1* as reference gene gave similar results; data not shown). Data are represented as median (25% percentile- 75% percentile).
[b] The median value of the non-infected group is set to 1.0.

In t-testing each gene is tested independently, the FOS vs cellulose effect was not statistically significant for each independent gene. However, FOS significantly increased expression of the cluster of the 26 Salmonella induced genes (see figure 5 and supplemental table 3.3). We also looked at overall gene expression differences between cellulose- and FOS-fed rats at day 2 after Salmonella infection (supplemental table 3.3). This was done to determine whether the stimulated translocation in FOS-fed rats resulted in additionally affected genes or biological processes not induced by Salmonella in the cellulose groups. Twenty genes were induced by Salmonella >2-fold

in cellulose-fed infected rats. In the FOS-fed infected rats 72 genes were induced by Salmonella >2-fold. Seventeen genes overlapped between these two diet groups. Detailed analysis of the genes exclusively induced (>2-fold) in the FOS-fed group showed that those could be categorized in the same processes identified earlier (table 2 and supplemental table 3.3). Obviously, the induced translocation of Salmonella by FOS supplementation did not affect other processes than those already identified in Salmonella infected rats on a cellulose diet. However, more genes of the same processes and higher fold-changes were noticed in the colonic mucosa of infected rats on the FOS diet.

Nos2 gene expression in colonic mucosa was below detection levels in the dietary infection study. As significant differences in NO_x excretion between infected cellulose- and FOS-fed rats were observed from day 5, no differences at *Nos2* gene expression were expected at day 2 p.i. Serum IFNγ was not detected at day 2 p.i., neither in infected cellulose-fed rats nor in infected FOS-fed rats.

Discussion

Colon is an infection target

This study shows quick and profound gene expression changes in the rat colon mucosa upon oral *S. enteritidis* infection, which implicates that not only the ileum, but also the colon, is a target for Salmonella infection. The earliest responses were noticed on mucosal transport and antimicrobial defense. The most responsive gene is *Pap*, which showed an 11-fold induction in colon mucosa on the first day after infection and increased to over 100 fold at day 6. At later timepoints, the most notable process affected is interferon-related. Colonic genes consistently induced by Salmonella infection in two independent studies, were all further enhanced by FOS supplementation, a known stimulus of colonic bacterial fermentation. Salmonella, ingested with contaminated foods or drinks, is thought to colonize the distal small intestine and to translocate through ileal Peyer's patches to extra-intestinal organs [48,49]. Several observations suggest that other parts of the intestine are also involved in Salmonella infection. High numbers of Salmonella are found in the cecum and colon of orally infected rats [48,49] as well as pigs [20]. In humans Salmonella commonly affects the small intestine, but colonic involvement of *S. enteritidis* has been reported in humans [41,46,60] and may play an important role in induction of diarrhea [41]. Studies describing mucosal invasion via the paracellular and transcellular route [27,38] also suggest that translocation of Salmonella species to the systemic circulation is not restricted to the ileal Peyer's patches. Our studies on the effects of prebiotics on resistance of the host to Salmonella infection also point to the colon as invasion site [5]. Together results from literature and those presented here indicate that the colon is one of the targets for Salmonella infection.

Interferon-gamma response

The increase of many IFNγ-regulated genes in the Salmonella-infected colon in the present *in vivo* study actually confirms the earlier suggested role of IFNγ in relation to host defense against Salmonella. Serum IFNγ levels increase in mice infected with Salmonella by oral or intraperitoneal route [9,14,31,32]. IFNγ is produced by natural killer cells, CD4 Th1 cells and CD8

cytotoxic lymphocytes [44,45,71]. IFNγ most likely exerts its function in host defense by activation of macrophages which can kill Salmonella [33]. In our Salmonella time course infection study, more than 20 IFNγ-related genes were up-regulated (table 2). This did not coincide with an increased IFNγ mRNA level at any of the time points studied. In addition, we could not detect IFNγ protein in Salmonella-infected colons (data not shown). Serum levels of this pro-inflammatory cytokine were undetectable at day 1 p.i. but rose steadily from day 3 p.i. with large inter-individual variation in the magnitude of response (figure 3). Despite the lack of detectable IFNγ protein in colonic mucosa and in serum at day 1, we did observe increased expression of genes in the IFNγ induced pathway at that timepoint. These genes are most likely activated by IFNγ [66]. We can not fully exclude that dilution of IFNγ-producing cells in the heterogeneous cell population of mucosal scrapings has lead to undetectable levels of this regulatory cytokine in the present study. At the later timepoints, serum IFNγ is strongly increased, whereas mucosal IFNγ remained below detection levels at all timepoints. This may suggest that systemic rather than colonic IFNγ seems to be the trigger for the later activation of IFNγ-related genes and -processes in colonic mucosa upon Salmonella infection. However, dilution of IFNγ-producing cells in colonic tissue to undetectable levels could also account for this later time point.

As many as 1200 genes are known to be regulated by IFNγ. Their gene products are mediators of the immune response essential for host defense against pathogens. One group of clearly regulated IFNγ-induced genes is the GTPase family, which modulates survival of pathogens residing in phagosomes or vacuoles [63,66]. They are defined into three classes: Guanylate-binding proteins (Gbp's), the p47 GTPases and the Mx proteins. We found Salmonella-induced up-regulation of the first and second group, i.e. *Gbp1, Ifi47, Gbp2, Iigp* and *Irgm* (Table 2). IFNγ induces expression of p47 GTPases via activation of Stat1 which was also increased by Salmonella at days 2 and 3 p.i. Mutant mice with gene disruptions in IFNγ or Stat1 are significantly compromised in their immune response to microbial infections, including salmonellosis [43]. Thus the increased expression of IFNγ-related genes in colonic mucosa in the present study confirms the earlier proposed role of this cytokine in Salmonella infection.

Pancreatits associated protein

The colonic mucosal gene most highly induced by Salmonella infection on the array was *Pap*, which was confirmed by Q-PCR of individual rat samples. *Pap* is a member of the *Reg III* gene family, which includes Regenerating islet-derived 3 gamma (*Reg3g*) which was also increased in our study. *Pap* expression is also increased in the rat ileal mucosa infected with *S. enteritidis* [58] and in the gastrointestinal tract of pigs infected with *Salmonella typimurium* [51]. Furthermore, significant up-regulation of intestinal mucosal *Pap* expression is described in IBD patients, whose bowel is chronically inflamed [12,19,52]. This suggests that PAP is a marker for acute as well as chronic inflammation. Biological functions of PAP in the intestine are not fully uncovered. Recently, it was proposed to function in innate immunity [8,19]. PAP was shown to have direct antimicrobial properties as it was able to bind and kill Gram-positive bacteria, but not Gram-negative *Salmonella typhimurium* [8]. Additional research will be needed to answer whether PAP is able to inhibit the growth of *Salmonella enteritidis*. *Pap* and *Reg3g* are expressed in several tissues and organs, but the small intestine has the highest expression under normal conditions. Only very low levels can be found in colon [29]. Indeed, *Pap* mRNA expression for most non-infected rats was below detection

level (figure 4). Three rats did not express *Pap* at timepoint day 3 p.i. (figure 4), whereas at day 6 p.i. all rats expressed increased levels of *Pap*. Variation in infection kinetics between (outbred) rats is obviously reflected in *Pap* expression. We are currently investigating which mucosal cell types contain PAP and whether it is secreted to the intestinal lumen or to the serosal (blood) site. If secreted, PAP might be used as a non-specific marker to follow and quantify intestinal infection or inflammation in humans.

Calprotectin

Calprotectin (*S100a8/a9*), a heterodimer of the two calcium-binding proteins S100A8 and S100A9, was up-regulated in the colonic mucosa by Salmonella (Table 2). Both subunits were increased in colon. Calprotectin is a 36 kDa calcium and zinc binding protein and constitutes approximately 60% of soluble cytosolic proteins in neutrophil granulocytes. Therefore, calprotectin is a marker of neutrophil influx and is elevated in a number of inflammatory conditions. In agreement with our results, Naughton et al (1996) also found increased levels of this marker in Salmonella- infected animals. Fecal calprotectin is emerging as a useful marker to quantify mucosal inflammation, not in the least because it appears to be stable in feces which can be obtained by non-invasive means [55].

Differences between colon and ileum

Ileum and colon are both targets for Salmonella. Remarkably, the number of genes showing increased expression is larger than the number of genes showing decreased expression upon Salmonella infection in both ileum and colon. However, this is more extreme in colon than in ileum mucosa [58]. Technical bias is unlikely as in a flavonoid intervention study with rats and using the same array system and data handling the number of down-regulated genes was similar to the number of up-regulated genes [11]. In an *in vivo* Salmonella infection study in pigs only up-regulated and no down-regulated genes were observed [51].

The extent of the early response to Salmonella is similar for both intestinal segments: From all genes expressed above background level on the arrays, 0.21% of the genes expressed in the colon and 0.26% of genes expressed in the ileum [58] were affected at day 1 p.i. The colonic response is less than the ileal response at day 3 p.i., as 0.15% of colonic mucosal genes were affected versus 0.67% of ileal mucosal genes. The smaller colonic response could be due to differences in crypt-villus architecture of the ileal and colonic mucosa. Furthermore, the colonic mucosa, which is constitutively exposed to bacteria, might be more efficient in repressing host- or more specifically immunological responses to bacteria, including pathogens [10,30,35].

Ileum and colon show overlapping as well as distinct processes affected upon oral infection [58]. At early time points after oral infection i.e. transport processes and antimicrobial defenses were regulated in both intestinal segments, but the process-related genes did not fully overlap. At day 1 p.i., glucose transporters were increased in the ileum, whereas in colon ion transporters were induced. The role of ion transporters in water absorption support involvement of the distal part of the gut in diarrhea development during salmonellosis as reported earlier in humans [41]. The gene coding for antimicrobial defensin 5 was down-regulated by Salmonella in both ileum and colon. Other genes coding for antimicrobial proteins (*Pla2g2a* and lysozym) were clearly enhanced in the infected colon in contrast to ileal tissue [58]. At day 3 and 6 p.i. Salmonella reduced

the expression of several phase I and II detoxification genes in the ileum, which was not observed in the colon. The downregulation of cytochrome P450 genes in ileum coincided with increased expression of inflammatory genes. It is known that inflammatory mediators can down-regulate cytochrome P450 genes[47,57,58]. This might suggest that the inflammatory response induced by Salmonella in colon, at later time-points, is smaller than in ileum. Nevertheless, both tissues showed signs of an inflammatory response at later time-points, but responsible genes were not the same. Mainly cytokines and chemokines were induced in the ileum, whereas in colon many interferon-related genes were up-regulated. No interferon response was observed in the ileum. Apparently, the immune response in the two intestinal segments is differentially regulated.

Finally, the *in vivo* transcriptional response of intact mucosa to invasion by Salmonella is represented by a limited number of regulated genes compared to *in vitro* studies with HT-29 cells[15]. *In vitro* models provide insight in complex mechanisms of Salmonella-host interaction. However, results should be interpreted with caution as *in vitro* systems show massive cell death at 24 hours, whereas only minor inflammatory changes are observed in the intestine 24 hours after infection with Salmonella *in vivo* [17]. Several genes like Toll like receptors, *Nf-κb* or *Il-8* that are regulated by Salmonella *in vitro*, were not found to be regulated by Salmonella infection in the present *in vivo* study. Possibly, transcription of these genes is highly specific for particular cell types in the colonic mucosa. Identification of cell type-specific responses of potential target cells could be addressed *in vivo* using laser microdissection.

FOS and mucosal barrier function

We consistently observed that diets supplemented with rapidly-fermentable prebiotics (such as FOS) increased translocation of *S. enteritidis* in rat infection studies despite stimulation of *Bifidobacteria* and *Lactobacilli* [68]. In other words, FOS decreases the resistance of the rat intestinal mucosa to intestinal pathogens. Because fermentation of FOS hardly occurs in the ileum of humans[2] and rats[22], it is unlikely that prebiotics facilitated translocation of Salmonella in the ileum. This is supported by the absence of ileal inflammation in FOS-fed and Salmonella-infected rats in contrast to profound cecal and colonic inflammation[5]. The precise mechanism underlying the effects of FOS on the colon mucosa is not known. FOS itself, the changed intestinal microflora or its fermentation products (e.g. SCFA) could play a role. Prebiotics, such as FOS resist enzymatic hydrolysis by digestive enzymes secreted in the small intestine and reach the colon intact. The resident colonic microflora ferments these carbohydrates to lactic acid and short-chain fatty acids (SCFA). This results in lowering of the pH of intestinal contents and stimulation of e.g. *Bifidobacteria* and *Lactobacilli*[67,68]. These lactic acid bacteria are assumed to enhance resistance [18] but we found opposite effects [5,68]. As shown earlier, dietary FOS increases intestinal permeability in non-infected rats and even more in infected rats[69]. At present it is unknown whether intestinal permeability is increased in ileum or colon, nor whether it is induced by the presence of FOS or by its fermentation metabolites. It has been shown that SCFA can induce colonic mucosal injury and increase permeability [3]. Furthermore, *in vitro* studies showed that SCFA can enhance expression of virulence (e.g. invasion) genes of *Salmonella typhimurium* [13,39], but data on *in vivo* consequences have not been reported. Preliminary experiments of our lab

showed no evidence for increased expression of virulence genes of *Salmonella enteritidis* in infected rats fed a FOS-diet (unpublished results).

From two independent rat infection studies we identified 26 colonic mucosal genes consistently affected by Salmonella. These 'robust' Salmonella target genes were all further induced by the FOS diet. The pronounced effects of FOS on Salmonella translocation were reflected by a modest but highly consistent increase of all Salmonella target genes. Moreover, the total number of genes induced by Salmonella is nearly 3 times higher in FOS-fed rats than in cellulose-fed rats. However, biological processes identified to be affected by Salmonella in colon of FOS-fed rats were not different from those observed in cellulose-fed rats. So, the quality of the colonic response was the same, but clearly the magnitude of the response was increased by FOS feeding. Based on the physiological effects, larger gene expression differences might have been expected. The modest responses observed might be due to our focus on Salmonella-induced gene expression. It can not be totally excluded that FOS targets other genes and processes related to barrier function (in absence of infection) than Salmonella. However, the genes affected by Salmonella in FOS-fed rats did not show involvement of additional processes in comparison to their cellulose-fed counterparts. In our view, the enhanced expression of colonic Salmonella target genes in FOS-fed animals concomitant with stimulated translocation of this invasive pathogen indicates that infection and related inflammation is worsened by FOS supplementation. Histological analyses of intestinal samples from previous FOS intervention studies of our lab did not show presence of intestinal mucosal inflammation in non-infected FOS-fed rats in contrast to post-infection samples (data not shown). Therefore, we feel that the observed aggravation of the intestinal response is due to interaction of FOS and Salmonella.

It should be stressed, that genes identified as Salmonella target genes in the present study are not necessarily Salmonella specific, but may well result from colonic inflammation in general and thus be similar in other enteric infections. Furthermore, effects of dietary FOS on gut barrier function may not be restricted to changes in mRNA expression, but exist on the translational or functional level of proteins. For instance, internalization of the tight-junction proteins occludin, claudin and junctional adhesion molecule-A, caused by IFNγ, results in profound mucosal barrier changes [7]. This cellular translocation can occur without concomitant changes in mRNA gene expression. Detection of such effects would require a different approach from transcriptomics. Many studies report on possible therapeutic effects of FOS on intestinal disease such as IBD and pathogenic infection. In addition to an increase in "beneficial" bacteria, the potential beneficial effects of FOS are based on the effects on surrogate markers, e.g. increase of mucin production [61], increase of the size and cytokine production of Peyer's patches and increased faecal or ileal IgA[25,59,65]. Changes in these markers are often presumed to reflect increased barrier function or resistance to pathogenic bacteria, but concomitant actual measurements of these functional effects are missing. In our study, genes involved in antimicrobial defense, immune response and inflammation were all induced by Salmonella infection and further enhanced by dietary FOS, but concomitantly translocation of Salmonella was evident and stimulated by FOS. Therefore, these surrogate markers should be interpreted with caution and always correlated with functional effects or clinical endpoints.

In this study we compared Salmonella-induced gene expression changes of two independent rat infection experiments at early time points after oral infection. The gene expression results were analyzed at two levels, at the level of gene expression itself and at the level of biological processes. Analysis at the level of gene expression showed some variation in the expression of individual genes between the two studies (table 2, supplemental table 3.3). This variation between two studies can be due to the different time-points studied and the use of outbred rats showing inter-individual differences in infection kinetics. Rats did not all respond to Salmonella at the same time p.i. which is e.g. shown by individual gene expression levels of PAP (figure 4) and by serum IFN levels (figure 3). Variation in infection kinetics and inter-individual variation are expected features of infection studies in outbred species [28]. It can be argued that differences between studies, due to differences in time points measured or infection kinetics, will result in more pronounced variance at the level of individual genes than at the level of physiological processes [62,64]. Indeed, analysis of gene expression at the level of biological processes showed that both studies gave highly comparable Salmonella-induced effects at early time points.

In the dietary infection study, we were interested whether the FOS-stimulated Salmonella translocation was reflected in colonic gene expression changes. We observed an overlap in gene expression changes observed in the two experimental diets, and an additional set of 58 genes which were only significantly affected in the FOS-fed rats. Although the list of altered genes was different in FOS-fed rats, this was not the case at the process level as exactly the same processes were observed for both dietary groups. This indicates involvement of similar underlying biological processes in cellulose and FOS-fed infected rats and no obvious role for other processes.

Comparison at the level of biological processes is a powerful tool to interpret microarray experiments and enables comparison of different microarray datasets[42]. Comparison at gene level has some drawbacks, one is redundancy in gene function, which means that different genes can provide the same physiological effect. In addition, the homeostatic condition as well as the precise nature of the stimulus will determine how individual genes within a process are controlled to provide the necessary physiological response. Differences in responses of individual genes are filtered out when they are analyzed at the level of pathways or processes. However, the translation of differentially expressed genes into biological processes also suffers from limitations [36]. The most important limitation is that annotations to pathways and processes are incomplete. Therefore it is important that results from pathway analysis are manually supplemented with the remaining significant genes using biological databases and scientific literature.

Conclusions

In conclusion, our results show that, in addition to the ileum, the colon mucosa is clearly a target for Salmonella infection. Early Salmonella-induced changes were observed in transport and oxidative stress, while at later stage, most likely secondary, infection and inflammation responses were observed. Some findings confirm expected results, such as induction of an immune and inflammatory response. However, the Salmonella-induced immune response in colon is clearly different from that in ileum. We newly identified that colonic transport processes and proteolysis

are affected by Salmonella infection and that pancreatitis associated protein was the most responsive gene in Salmonella infected rat colon.

An important observation is that FOS-stimulated Salmonella translocation (as measured by urinary NO_x), does not induce other processes than those observed in cellulose-fed and Salmonella infected rats. So, the quality or diversity of the colonic host response to Salmonella is not affected by colonic FOS fermentation in contrast to the magnitude of response. As far as we know, there are no literature data pointing to a functional effect of FOS in the ileum. Therefore, the FOS effects on Salmonella translocation are most likely due to colonic effects. Understanding the changes caused by FOS alone may provide insight in processes that ultimately result in the observed weakening of the barrier.

Acknowledgements

References

1. Anderle, P., Rakhmanova, V., Woodford, K., Zerangue, N. & Sadee, W. Messenger RNA expression of transporter and ion channel genes in undifferentiated and differentiated Caco-2 cells compared to human intestines. *Pharm Res* 20, 3-15 (2003).
2. Andersson, H.B., Ellegard, L.H. & Bosaeus, I.G. Nondigestibility characteristics of inulin and oligofructose in humans. *J Nutr* 129, 1428S-1430S (1999).
3. Argenzio, R.A. & Meuten, D.J. Short-chain fatty acids induce reversible injury of porcine colon. *Dig Dis Sci* 36, 1459-1468 (1991).
4. Barrow, P.A. Salmonella infections: immune and non-immune protection with vaccines. *Avian Pathol* 36, 1-13 (2007).
5. Bovee-Oudenhoven, I.M., ten Bruggencate, S.J., Lettink-Wissink, M.L. & van der Meer, R. Dietary fructo-oligosaccharides and lactulose inhibit intestinal colonisation but stimulate translocation of salmonella in rats. *Gut* 52, 1572-1578 (2003).
6. Bovee-Oudenhoven, I.M., Termont, D.S., Weerkamp, A.H., Faassen-Peters, M.A. & Van der Meer, R. Dietary calcium inhibits the intestinal colonization and translocation of Salmonella in rats. *Gastroenterology* 113, 550-557 (1997).
7. Bruewer, M*., et al.* Interferon-gamma induces internalization of epithelial tight junction proteins via a macropinocytosis-like process. *Faseb J* 19, 923-933 (2005).
8. Cash, H.L., Whitham, C.V., Behrendt, C.L. & Hooper, L.V. Symbiotic bacteria direct expression of an intestinal bactericidal lectin. *Science* 313, 1126-1130 (2006).
9. Chong, C., Bost, K.L. & Clements, J.D. Differential production of interleukin-12 mRNA by murine macrophages in response to viable or killed Salmonella spp. *Infect Immun* 64, 1154-1160 (1996).
10. Collier-Hyams, L.S., Sloane, V., Batten, B.C. & Neish, A.S. Cutting edge: bacterial modulation of epithelial signaling via changes in neddylation of cullin-1. *J Immunol* 175, 4194-4198 (2005).
11. de Boer, V.C*., et al.* Chronic quercetin exposure affects fatty acid catabolism in rat lung. *Cell Mol Life Sci* 63, 2847-2858 (2006).
12. Dieckgraefe, B.K., Stenson, W.F., Korzenik, J.R., Swanson, P.E. & Harrington, C.A. Analysis of mucosal gene expression in inflammatory bowel disease by parallel oligonucleotide arrays. *Physiol Genomics* 4, 1-11 (2000).

13. Durant, J.A., Corrier, D.E. & Ricke, S.C. Short-chain volatile fatty acids modulate the expression of the hilA and invF genes of Salmonella typhimurium. *J Food Prot* 63, 573-578 (2000).

14. Eckmann, L. & Kagnoff, M.F. Cytokines in host defense against Salmonella. *Microbes Infect* 3, 1191-1200 (2001).

15. Eckmann, L., Smith, J.R., Housley, M.P., Dwinell, M.B. & Kagnoff, M.F. Analysis by high density cDNA arrays of altered gene expression in human intestinal epithelial cells in response to infection with the invasive enteric bacteria Salmonella. *J Biol Chem* 275, 14084-14094 (2000).

16. Ekins, S., Nikolsky, Y., Bugrim, A., Kirillov, E. & Nikolskaya, T. Pathway mapping tools for analysis of high content data. *Methods Mol Biol* 356, 319-350 (2007).

17. Finlay, B.B. & Brumell, J.H. Salmonella interactions with host cells: in vitro to in vivo. *Philos Trans R Soc Lond B Biol Sci* 355, 623-631 (2000).

18. Gibson, G.R. Dietary modulation of the human gut microflora using the prebiotics oligofructose and inulin. *J Nutr* 129, 1438S-1441S (1999).

19. Gironella, M.*, et al.* Anti-inflammatory effects of pancreatitis associated protein in inflammatory bowel disease. *Gut* 54, 1244-1253 (2005).

20. Gray, J.T., Fedorka-Cray, P.J., Stabel, T.J. & Ackermann, M.R. Influence of inoculation route on the carrier state of Salmonella choleraesuis in swine. *Vet Microbiol* 47, 43-59 (1995).

21. Havelaar, A.H.*, et al.* A rat model for dose-response relationships of Salmonella Enteritidis infection. *J Appl Microbiol* 91, 442-452 (2001).

22. Heijnen, A.M., Brink, E.J., Lemmens, A.G. & Beynen, A.C. Ileal pH and apparent absorption of magnesium in rats fed on diets containing either lactose or lactulose. *Br J Nutr* 70, 747-756 (1993).

23. Herikstad, H., Motarjemi, Y. & Tauxe, R.V. Salmonella surveillance: a global survey of public health serotyping. *Epidemiol Infect* 129, 1-8 (2002).

24. Honer zu Bentrup, K.*, et al.* Three-dimensional organotypic models of human colonic epithelium to study the early stages of enteric salmonellosis. *Microbes Infect* 8, 1813-1825 (2006).

25. Hosono, A.*, et al.* Dietary fructooligosaccharides induce immunoregulation of intestinal IgA secretion by murine Peyer's patch cells. *Biosci Biotechnol Biochem* 67, 758-764 (2003).

26. Hsiao, L.L.*, et al.* A compendium of gene expression in normal human tissues. *Physiol Genomics* 7, 97-104 (2001).

27. Hughes, E.A. & Galan, J.E. Immune response to Salmonella: location, location, location? *Immunity* 16, 325-328 (2002).

28. Hyland, K.A., Kohrt, L., Vulchanova, L. & Murtaugh, M.P. Mucosal innate immune response to intragastric infection by Salmonella enterica serovar Choleraesuis. *Mol Immunol* 43, 1890-1899 (2006).

29. Iovanna, J.L.*, et al.* PAP, a pancreatic secretory protein induced during acute pancreatitis, is expressed in rat intestine. *Am J Physiol* 265, G611-618 (1993).

30. Ismail, A.S. & Hooper, L.V. Epithelial cells and their neighbors. IV. Bacterial contributions to intestinal epithelial barrier integrity. *Am J Physiol Gastrointest Liver Physiol* 289, G779-784 (2005).

31. John, B.*, et al.* Role of IL-12-independent and IL-12-dependent pathways in regulating generation of the IFN-gamma component of T cell responses to Salmonella typhimurium. *J Immunol* 169, 2545-2552 (2002).

32. Jouanguy, E.*, et al.* IL-12 and IFN-gamma in host defense against mycobacteria and salmonella in mice and men. *Curr Opin Immunol* 11, 346-351 (1999).

33. Kagaya, K., Watanabe, K. & Fukazawa, Y. Capacity of recombinant gamma interferon to activate macrophages for Salmonella-killing activity. *Infect Immun* 57, 609-615 (1989).

34. Kagnoff, M.F. & Eckmann, L. Epithelial cells as sensors for microbial infection. *J Clin Invest* 100, 6-10 (1997).

35. Kelly, D.*, et al.* Commensal anaerobic gut bacteria attenuate inflammation by regulating nuclear-cytoplasmic shuttling of PPAR-gamma and RelA. *Nat Immunol* 5, 104-112 (2004).

36. Khatri, P. & Draghici, S. Ontological analysis of gene expression data: current tools, limitations, and open problems. *Bioinformatics* 21, 3587-3595 (2005).

37. Knight, P.A.*, et al.* Expression profiling reveals novel innate and inflammatory responses in the jejunal epithelial compartment during infection with Trichinella spiralis. *Infect Immun* 72, 6076-6086 (2004).

38. Kops, S.K., Lowe, D.K., Bement, W.M. & West, A.B. Migration of Salmonella typhi through intestinal epithelial monolayers: an in vitro study. *Microbiol Immunol* 40, 799-811 (1996).

39. Lawhon, S.D., Maurer, R., Suyemoto, M. & Altier, C. Intestinal short-chain fatty acids alter Salmonella typhimurium invasion gene expression and virulence through BarA/SirA. *Mol Microbiol* 46, 1451-1464 (2002).

40.	Lee, H.K., Braynen, W., Keshav, K. & Pavlidis, P. ErmineJ: tool for functional analysis of gene expression data sets. *BMC Bioinformatics* 6, 269 (2005).

41.	Mandal, B.K. & Mani, V. Colonic involvement in salmonellosis. *Lancet* 1, 887-888 (1976).

42.	Manoli, T.*, et al.* Group testing for pathway analysis improves comparability of different microarray datasets. *Bioinformatics* 22, 2500-2506 (2006).

43.	Mastroeni, P.*, et al.* Interleukin 18 contributes to host resistance and gamma interferon production in mice infected with virulent Salmonella typhimurium. *Infect Immun* 67, 478-483 (1999).

44.	Mastroeni, P., Harrison, J.A., Chabalgoity, J.A. & Hormaeche, C.E. Effect of interleukin 12 neutralization on host resistance and gamma interferon production in mouse typhoid. *Infect Immun* 64, 189-196 (1996).

45.	Mastroeni, P.*, et al.* Interleukin-12 is required for control of the growth of attenuated aromatic-compound-dependent salmonellae in BALB/c mice: role of gamma interferon and macrophage activation. *Infect Immun* 66, 4767-4776 (1998).

46.	McGovern, V.J. & Slavutin, L.J. Pathology of salmonella colitis. *Am J Surg Pathol* 3, 483-490 (1979).

47.	Morgan, E.T., Li-Masters, T. & Cheng, P.Y. Mechanisms of cytochrome P450 regulation by inflammatory mediators. *Toxicology* 181-182, 207-210 (2002).

48.	Naughton, P.J.*, et al.* Salmonella typhimurium and Salmonella enteritidis induce gut growth and increase the polyamine content of the rat small intestine in vivo. *FEMS Immunol Med Microbiol* 12, 251-258 (1995).

49.	Naughton, P.J., Grant, G., Spencer, R.J., Bardocz, S. & Pusztai, A. A rat model of infection by Salmonella typhimurium or Salm. enteritidis. *J Appl Bacteriol* 81, 651-656 (1996).

50.	NCBI GEO website. http://www.ncbi.nlm.nih.gov/geo/. (,).

51.	Niewold, T.A.*, et al.* The early transcriptional response of pig small intestinal mucosa to invasion by Salmonella enterica serovar typhimurium DT104. *Mol Immunol* 44, 1316-1322 (2007).

52.	Ogawa, H.*, et al.* Increased expression of HIP/PAP and regenerating gene III in human inflammatory bowel disease and a murine bacterial reconstitution model. *Inflamm Bowel Dis* 9, 162-170 (2003).

53.	Oudenhoven, I.M., Klaasen, H.L., Lapre, J.A., Weerkamp, A.H. & Van der Meer, R. Nitric oxide-derived urinary nitrate as a marker of intestinal bacterial translocation in rats. *Gastroenterology* 107, 47-53 (1994).

54.	Pellis, L., Franssen-van Hal, N.L., Burema, J. & Keijer, J. The intraclass correlation coefficient applied for evaluation of data correction, labeling methods, and rectal biopsy sampling in DNA microarray experiments. *Physiol Genomics* 16, 99-106 (2003).

55.	Poullis, A., Foster, R., Northfield, T.C. & Mendall, M.A. Review article: faecal markers in the assessment of activity in inflammatory bowel disease. *Aliment Pharmacol Ther* 16, 675-681 (2002).

56.	Reeves, P.G., Nielsen, F.H. & Fahey, G.C., Jr. AIN-93 purified diets for laboratory rodents: final report of the American Institute of Nutrition ad hoc writing committee on the reformulation of the AIN-76A rodent diet. *J Nutr* 123, 1939-1951 (1993).

57.	Renton, K.W. Alteration of drug biotransformation and elimination during infection and inflammation. *Pharmacol Ther* 92, 147-163 (2001).

58.	Rodenburg, W., Bovee-Oudenhoven, I.M.J., Kramer, E., van der Meer, R. & Keijer, J. Gene expression response of the rat small intestine following oral Salmonella infection. *Physiological Genomics* 30, 123-133 (2007).

59.	Roller, M., Rechkemmer, G. & Watzl, B. Prebiotic inulin enriched with oligofructose in combination with the probiotics Lactobacillus rhamnosus and Bifidobacterium lactis modulates intestinal immune functions in rats. *J Nutr* 134, 153-156 (2004).

60.	Rout, W.R., Formal, S.B., Dammin, G.J. & Giannella, R.A. Pathophysiology of Salmonella diarrhea in the Rhesus monkey: Intestinal transport, morphological and bacteriological studies. *Gastroenterology* 67, 59-70 (1974).

61.	Schmidt-Wittig, U., Enss, M.L., Coenen, M., Gartner, K. & Hedrich, H.J. Response of rat colonic mucosa to a high fiber diet. *Ann Nutr Metab* 40, 343-350 (1996).

62.	Segal, E., Friedman, N., Kaminski, N., Regev, A. & Koller, D. From signatures to models: understanding cancer using microarrays. *Nat Genet* 37 Suppl, S38-45 (2005).

63.	Singh, S.B., Davis, A.S., Taylor, G.A. & Deretic, V. Human IRGM induces autophagy to eliminate intracellular mycobacteria. *Science* 313, 1438-1441 (2006).

64.	Subramanian, A.*, et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102, 15545-15550 (2005).

65.	Swanson, K.S.*, et al.* Supplemental fructooligosaccharides and mannanoligosaccharides influence immune function, ileal and total tract nutrient digestibilities, microbial populations and concentrations of protein catabolites in the large bowel of dogs. *J Nutr* 132, 980-989 (2002).

66.   Taylor, G.A., Feng, C.G. & Sher, A. p47 GTPases: regulators of immunity to intracellular pathogens. *Nat Rev Immunol* 4, 100-109 (2004).

67.   Ten Bruggencate, S.J., Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L., Katan, M.B. & van der Meer, R. Dietary fructooligosaccharides affect intestinal barrier function in healthy men. *J Nutr* 136, 70-74 (2006).

68.   Ten Bruggencate, S.J., Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L. & Van der Meer, R. Dietary fructo-oligosaccharides dose-dependently increase translocation of salmonella in rats. *J Nutr* 133, 2313-2318 (2003).

69.   Ten Bruggencate, S.J., Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L. & Van der Meer, R. Dietary fructooligosaccharides increase intestinal permeability in rats. *J Nutr* 135, 837-842 (2005).

70.   van Hal, N.L.*, et al.* The application of DNA microarrays in gene expression analysis. *J Biotechnol* 78, 271-280 (2000).

71.   VanCott, J.L.*, et al.* Regulation of mucosal and systemic antibody responses by T helper cell subsets, macrophages, and derived cytokines following oral immunization with live recombinant Salmonella. *J Immunol* 156, 1504-1514 (1996).

72.   Vender, R.J. & Marignani, P. Salmonella colitis presenting as a segmental colitis resembling Crohn's disease. *Dig Dis Sci* 28, 848-851 (1983).

73.   Wettenhall, J.M. & Smyth, G.K. limmaGUI: a graphical user interface for linear modeling of microarray data. *Bioinformatics* 20, 3705-3706 (2004).

Chapter 4

Ileal mucosal and faecal pancreatitis associated protein is increased during Salmonella infection in rats and is associated with infection severity

Marleen T. J. van Ampting [1, 2]
Wendy Rodenburg[1, 2, 3]
Carolien Vink[1, 2]
Evelien Kramer[1, 3]
Arjan J. Schonewille[1, 2]
Jaap Keijer[1, 3]
Roelof van der Meer[1, 2, 4]
Ingeborg M.J. Bovee-Oudenhoven[1, 2]

[1] TI Food and Nutrition, Wageningen, The Netherlands
[3] RIKILT Institute of Food Safety, Wageningen, The Netherlands
[2] NIZO Food Research, Ede, The Netherlands
[4] Nutrigenomics Consortium, TI Food and Nutrition, Wageningen, The Netherlands

ABSTRACT

Background

Microbial infection and increased microbial-epithelial contact induce ileal pancreatitis associated protein/regenerating gene III (*PAP/RegIII*) mRNA expression. PAP/RegIII is suggested to function as a stress protein with anti-inflammatory or antimicrobial functions. Up to now in vivo PAP/RegIII protein expression has been poorly studied. Therefore we aimed to study PAP/RegIII protein levels in the ileal mucosa and faeces of rats infected with Salmonella. Rats were fed a diet relatively low or high in calcium to decrease or improve intestinal resistance to infection, respectively.

Design

Rats on a Western-type diet containing 30 or 120 mmol/kg calcium were orally infected with *Salmonella enteritidis*. At days 3-4 post infection, the ileal mucosa was isolated to determine PAP/RegIII mRNA and protein expression. Parallel groups were studied until days 7-8 post-infection to determine Salmonella colonisation and translocation and to assess PAP/RegIII excretion in faeces with time.

Results

Salmonella infection significantly increased ileal mucosal PAP/RegIII expression on mRNA and protein level in comparison with non-infected controls. Immunohistology showed that PAP/RegIII was present in epithelial cells located at the crypt-villus junction. Faecal PAP/RegIII excretion increased after infection. Inhibition of Salmonella colonisation and translocation by dietary calcium was associated with lower mucosal and faecal PAP/RegIII concentrations.

Conclusion

PAP/RegIII protein is increased in epithelial cells of the ileal mucosa during Salmonella infection and is associated with infection severity. This supports its proposed role as an intestinal stress protein. PAP/RegIII is excreted in faeces and might be used as infection marker.

Introduction

Pancreatitis associated protein (PAP) is a type III member of the regenerating (Reg) gene family and was originally identified as a lectin-related secretory protein present in rat pancreatic juice during experimental pancreatitis[19]. Since then, considerable attention has been given to the Reg family and its structurally related molecules. Recently, the complex terminology of the Reg family and its isoforms was elegantly reviewed[15] and a combined term of PAP/RegIII was coined, to foster a concerted effort in the investigation of PAP and the isoforms. In this study we focused on ileal PAP/RegIII, represented by the genes *PAP1* and *PAP3*. Expression of *PAP/RegIII* mRNA has been shown in the pancreas of human,[37] mouse[35] and rat[17]. Moreover, *PAP/RegIII* was described as constitutively expressed in the rat[18] and human[9] small intestine and this expression was not altered during acute pancreatitis.[18] Interestingly, increased levels of intestinal *PAP/RegIII* mRNA have been detected during active inflammatory bowel disease (IBD), Crohn's disease and ulcerative colitis, in humans[11,23] and in animal models of IBD[14,23,32]. We recently reported a time dependent increase in intestinal *PAP/RegIII* mRNA, represented by the genes *PAP1* and *PAP3,* in Salmonella infection studies in rats.[27,28] These results are supported by studies that showed increased *PAP/RegIII* mRNA upon bacterial colonisation of the porcine small intestine with *Salmonella*[22] and enterotoxigenic *Escherichia coli*.[21] Hence, it is suggested that PAP/RegIII expression is triggered by increased microbial-epithelial contact and reflects a state of enhanced host defence[8]. Hitherto, few studies focus on PAP/RegIII protein levels and it was not investigated whether a more severe infection or increased host defence status would result in a higher PAP/RegIII expression.

Furthermore, data about the site of ileal PAP/RegIII expression is controversial. *PAP/RegIII* mRNA is reported to be present in epithelial cells of the lower villus part,[18] other studies limit expression to Paneth cells[8].

Although several functions have been proposed for PAP/RegIII, the physiological relevance of *PAP/RegIII* mRNA upregulation under conditions of infection and inflammation is unknown at present. In a search of a functional role for PAP/RegIII, the mouse RegIIIγ isoform showed in vitro binding to carbohydrates present on the surface of bacteria[8]. Moreover, mouse and human RegIIIγ were shown to have anti-bacterial activity against Gram-positive bacteria.[8] Other studies suggested that PAP/RegIII may have anti-inflammatory functions by blocking NF-κB activation[12,14]. These anti-inflammatory effects were shown for experimental pancreatitis[36,38], but whether the same function can be extrapolated to the intestine remains to be explored.

To investigate intestinal localisation and protein levels of PAP/RegIII we investigated PAP/RegIII expression in the rat small intestine upon infection with *Salmonella enteritidis*, which is a common foodborne pathogen. We focussed our study on the distal ileum, since this was shown to be the main site of PAP/RegIII expression (genes *PAP1* and *PAP3*) upon Salmonella infection[27,28]. In addition, we assessed whether ileal PAP/RegIII expression is sensitive to dietary modulation of infection severity. We have shown that dietary calcium has profound resistance-enhancing effects and protects against infectious diarrhoea in rats[4,5] and humans[2]. Calcium supplementation inhibits colonisation and translocation of invasive *S. enteritidis*.[5] Studying rat intestinal PAP/RegIII expression in a dietary calcium intervention and correlating its protein levels with well-established markers of infection severity may extend insight in PAP/RegIII

functionality in the gut. Moreover, we determined its presence in faeces and assessed its proposed anti-microbial function.

Materials and Methods

Diets, infection and dissection of the rats

The experimental protocol was approved by the animal welfare committee of Wageningen University (Wageningen, the Netherlands). Specific pathogen-free male outbred Wistar rats (WU, Harlan, Horst, the Netherlands), 8 weeks old and with a mean body weight of 245 g, were housed individually in metabolic cages as described.[3] Rats were fed purified diets containing per kg: 200 g acid casein, 326 g cornstarch, 174 g glucose, 160 g palm oil, 40 g corn oil, 50 g cellulose and vitamin and mineral mix (without calcium) according to AIN-93.[26] To mimic the composition of a Western human diet, the prepared diets were relatively low in calcium and high in fat content in comparison with recommendations for rodent diets of the AIN-93. Diets were supplemented with $CaHPO_4.2H_2O$ (Merck, Darmstadt, Germany), at the expense of glucose, to a final concentration of 30 mmol/kg (control diet) or 120 mmol/kg (calcium diet). Food, intake recorded daily, and demineralised drinking water were supplied ad libitum. Body weight was measured every two days before infection and daily after infection. Two groups were fed the control diet and another two groups were fed the calcium-supplemented diet (n=9 per group). In addition, five rats were fed the control (n=3) or calcium (n=2) diet and served as non-infected controls.

Animals were acclimatised to housing and dietary conditions for 14 days, after which they were orally infected with 0.5 ml of saline containing 3.10^9 colony-forming units of *S. enteritidis* (clinical isolate, phage type 4; strain B1214 NIZO food research, Ede, the Netherlands) as described elsewhere.[3] Animals in the non-infected group received orally 0.5 ml saline only.

On day 3 or 4 after oral infection, one control and one calcium-supplemented group were randomly selected and killed by carbon dioxide inhalation. The non-infected control rats were also killed by this procedure. During the dissection, the distal 12 cm of the ileum was excised. From this a 2 cm piece, identical location in each sample, was cut out and preserved in 10% formalin (Sigma-Aldrich, St Louis, Missouri, USA) and embedded in paraffin for histological analysis. The remaining parts were cut open longitudinally and, after flushing with saline, the mucosa was scraped off and immediately frozen in liquid nitrogen for RNA isolation and protein analyses.

The other infected control and calcium groups were followed until day 7 after infection to collect fresh faecal samples for Salmonella quantification, as described elsewhere.[6] In addition, 24 h faeces (pooled per animal per 2 days) and urines were collected one day before infection and seven consecutive days after oral infection. All faeces and urine were stored at -20°C until further analysis. Oxytetracycline (Sigma-Aldrich) was added to the urine collection vessels of the metabolic cages to prevent bacterial deterioration. Bacterial translocation was quantified by measuring urinary NO_x (sum of nitrate and nitrite) excretion by using a colorimetric enzymatic kit (Roche Diagnostics, Basel, Switzerland), as described elsewhere[33].

Myeloperoxidase analysis in ileal mucosa

Frozen mucosal scrapings of the ileum were pulverized under liquid nitrogen. Approximately half of the pulverized tissue was suspended in a 0.2 M sucrose buffer of pH 7.4 containing 20 mM trishydroxymethylaminomethane (Tris), 1 mM dithiothreitol (DTT) and Complete Protease Inhibitor Cocktail (Roche Diagnostics). After mixing and centrifugation at 14000 g for 20 min the pellet was resuspended in acetate-HETAB buffer (0.5% hexadecyltrimethylammonium bromide (HETAB) at pH 6.0, 50 mM sodium acetate, 10 mM ethylenediaminetetraacetic acid (EDTA) and 0.25 M sucrose) and sonicated on ice for 30 s at level 2-3 (Sonicator XL2020, Heat Systems, Farmingdale, NY, USA). The protein concentration of the samples was determined using BC Assay (Omnilabo, the Netherlands) according to the manufacturer's protocol. A mouse myeloperoxidase (MPO) ELISA test kit (Hycult biotechnology, Uden, the Netherlands), which is cross-reactive with rat MPO, was used according to the manufacturer's guidelines to determine the concentration of MPO in mucosal scrapings.

Quantitative real-time PCR analysis of PAP/RegIII, represented by *PAP1* and *PAP3* mRNA, in ileal mucosa

The other half of the pulverized ileal mucosal scrapings was dissolved in TRIzol reagent (Invitrogen, Carlsbad, California, USA) to isolate and purify total RNA as described before.[27] By using TaqMan Reverse Transcription reagents (Applied Biosystems Inc., Foster City, California, USA) cDNA was created from 1 μg of RNA on a Perkin Elmer DNA Thermal Cycler 480, followed by SYBR Green-based real-time PCR on a 7500 Fast Real-Time PCR system (Applied Biosystems). PCR conditions used were 95 °C for 10 min, followed by 40 amplification cycles (95°C for 15 s, 60°C for 1 min). Data were normalised against β-actin and aldolase. Controls, methods and primerdesign were performed as described[27]. The primer sequences are listed in table 1.

Table 1. Sequences of primers used for quantitative real-time PCR analysis.

Gene	Acc nr	forward primer (5'→3')	reverse primer (5'→3')
PAP1	NM_053289	GACTCCATGACCCCACTCTTG	GCAGACGTAGGGCAACTT CAC
PAP3	NM_173097	GCTTCCTTTGTGTCCTCCTTGATT	TACTCCACTCCCATCCACCTCTG
β-actin	NM_031144	CTTTCTACAATGAGCTGCGTGTG	GTCAGGATCTTCATGAGGTAGTCTGTC
aldolase	NM_012495	ATGCCCCACCCATACCCAGCACT	AGCAGCAGTTGGCGGTAGAAGCG

Analysis of PAP/RegIII protein in ileal mucosa

Ileum paraffin sections were immunostained with a goat polyclonal antibody against rat PAP/RegIII (1:50) (PAP/ RegIII #AF 1996; R&D Systems, Minneapolis, MN, USA). This antibody can not discriminate between PAP1 and RegIII (PAP3), but since both *PAP1* and *PAP3* fall within the PAP/RegIII group, this discrimination is not essential for our investigation.[15] Localisation of Paneth cells was confirmed by immunostaining for lysozyme with ready to use anti-lysozyme (N1515; Dako). DAB peroxidase substrate kit (Dako) was used for signal detection of the HRP labelled secondary antibody, according to the manufacturer's protocol.

PAP/RegIII protein expression in the ileal mucosa was semi-quantified by light microscopy as follows: total length of (bottom) crypt to villus (tip) and part of this length stained positive for

PAP/RegIII were measured. The length stained positive for PAP/RegIII was expressed as percentage of total crypt-villus length. For each tissue section three completely visible crypt-villus axes were analysed, scored and averaged for that particular rat. All histological slides were recoded before microscopy to ensure blind scoring and to prevent observer's bias.

Analysis of PAP/RegIII protein in faeces

Total 24 h faeces were lyophilized in a manifold freeze dryer (FD5515; Ilshin Laboratory Co Ltd, Seoul, South Korea) and pooled per treatment group on the basis of individual daily faecal dry weight excretion. Proteins were isolated from lyophilized faeces pools as described elsewhere[16] with a few modifications. Briefly, 100 mg of faeces pool was homogenised in 500 µl buffer containing 50 mmol/L Tris-HCl (pH7.5), 100 mmol/L NaCl, 1 mmol/L EDTA and Complete Protease Inhibitor Cocktail (Roche). After centrifugation (2 min at 15000 g), the supernatant was taken and its protein concentration was determined using DC protein assay kit (Bio-rad Laboratories, Veenendaal, the Netherlands) according to the manufacturer's protocol. 45 µg protein was denatured at 100°C for 3 min in Tricine sample buffer (Bio-rad Laboratories), subjected to SDS-PAGE gel (4% stacking-gel, 14% separation-gel) and transferred to a PVDF membrane (Bio-rad Laboratories). After blocking, the membranes were incubated with the PAP/RegIII antibody (1:100). The signal of the secondary HRP-conjugated antibody was detected using the ECL Plus chemiluminescent detection kit (GE Healthcare, Den Bosch, the Netherlands).

To evaluate recovery of PAP/RegIII protein in faeces, 3x 50 mg of the pre-infection faeces pool, from animals fed the control diet, was homogenized in PBS. Of these three samples one was spiked with 1 µg recombinant rat PAP (recPAP; pre-release reagent from R&D Systems Inc.). The second sample was heat inactivated by incubation at 75°C for 10 min, then cooled down to room temperature and identically spiked with the recombinant protein. Just PBS was added to the third sample and used as negative control. Subsequently, the three samples were incubated at 37°C for 1 h and protein was isolated and analysed by immunoblotting as described above.

To assess in vitro bactericidal activity of PAP/RegIII against *S. enteritidis* and *Listeria monocytogenes*, analysis was performed as described[8] with few modifications. To mimic the natural environment were PAP/RegIII could function as antimicrobial peptide, we performed the assays in (sterile) faecal water extracts from non-infected animals fed the control diet, prepared as described.[34] Purified recPAP was added to incubates to a final concentration of 3.3 µM and viable pathogens were quantified after 0, 2, 4, 8 and 24 hours of incubation.

Statistical analysis

Data from the non-infected rats fed the control or calcium diet were pooled as no diet-induced differences were observed. All data are expressed as means ± SE, except for PCR results, which are individually plotted in addition to indication of groups' mean. Data were tested for normality by the Kolmogorov-Smirnov test. If normally distributed, differences of the means were tested for significance using one-way ANOVA, followed by Student's t-test (two sided). For non-normally distributed data, differences between means were tested for their significance using Kruskall-Wallis ANOVA, followed by the non-parametric Mann-Whitney U test (two sided). Statistical significance was set at $p < 0.05$.

Results

Animals and food intake

All data from one animal in the calcium group (followed until day 7-8 after infection) were excluded from the study results because that rat suffered from pneumonia due to oropharyngeal reflux of the S. *enteritidis* suspension. At the start of the experiment, mean body weight of the animals was 245 g. Average body weight gain (mean before infection 5 g/d, after infection 3.3 g/d) and food intake (mean 17 g dry wt/d; not affected by Salmonella infection) were not affected by dietary treatment (data not shown).

Faecal excretion and translocation of Salmonella

As expected, no Salmonella could be detected in faeces collected before infection of the animals. The first days after Salmonella infection, rats fed the calcium diet had approximately 10-fold less Salmonella in their faeces than rats fed the control diet (fig 1A; $p<0.05$). This result indicates much better colonisation resistance in rats fed the calcium diet.[6] The impaired colonisation resistance of rats fed the control diet coincided with a progressive increase in urinary NO_x excretion (fig 1B), which was shown to reflect differences in magnitude of bacterial translocation.[5] Total infection-induced urinary NO_x excretion (area under the curve) of the calcium group was significantly lower than that of the control group: 25 ± 5 versus 73 ± 14 μmol per 7 days, respectively.
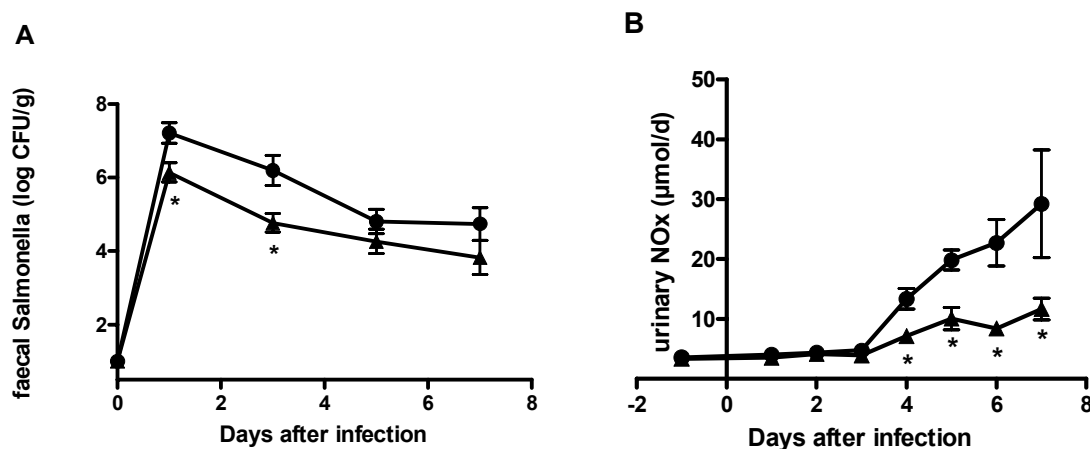
Figure 1. Effect of dietary calcium on (A) faecal Salmonella excretion and (B) infection induced urinary NO_x (sum of nitrate and nitrite) excretion with time. Rats fed the control ($\bullet$) or calcium ($\blacktriangle$) diet were orally infected with 3.10^9 colony-forming units *S. enteritidis* on day zero. Faeces and urine collected on day zero was prior to infection. Results are expressed as means +/- SE. An asterisk indicates a significant difference from the control group ($p<0.05$).

Ileal mucosal inflammation

On days 3-4 after Salmonella infection, compared to non-infected animals MPO levels increased by factor 11.5 and 5.4 in the control and calcium group, respectively (fig 2). Although post-infection MPO levels in the calcium-supplemented rats were half of those detected in the control group, this difference did not reach statistical significance ($p=0.3$). This was likely due to the relatively large inter-individual variation observed in the Salmonella-infected control group.
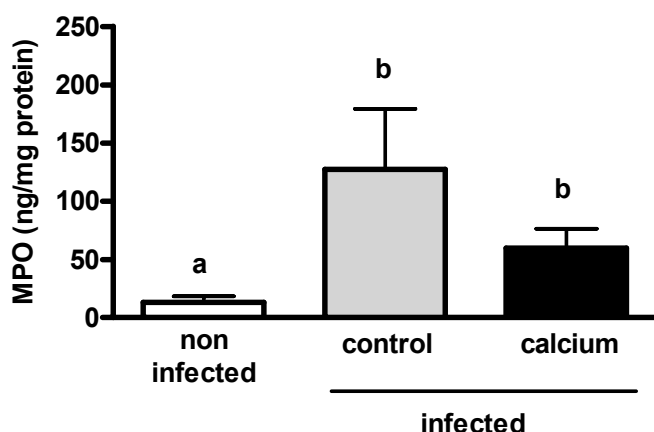
Figure 2. Myeloperoxidase (MPO) levels in the ileum mucosa of non-infected (n=5) and Salmonella infected rats fed either the control (n=9) or calcium-supplemented (n=9) diet. Results are expressed as means ± SE. Different letters indicate significant differences (p<0.05). Standard additions with MPO, provided by the manufacturer, were included in each assay to confirm the absence of inhibiting factors in the ileal mucosal samples.

Effect of infection and dietary calcium intervention on PAP/RegIII mRNA and protein expression in ileal mucosa

During infection ileal *PAP1* is 3 fold higher expressed than *PAP3* (*PAP1*/*actin* was 0.8 and 0.7 and *PAP3*/*Actin* was 0.29 and 0.27 in control and calcium groups, respectively). Salmonella infection of rats fed the control diet increased *PAP1* mRNA 2.4-fold, whereas no up-regulation was observed in calcium-supplemented infected rats (p<0.05, fig 3A). Furthermore, compared to non-infected rats a 3.6-and 1.8-fold induction of *PAP3* mRNA was observed in infected rats fed the control and calcium diet, respectively (fig 3B).

PAP/RegIII protein was detected in the ileum mucosa of non-infected rats (fig 4). Infected animals showed an increased number of cells staining positive for PAP/RegIII protein (fig 4). PAP/RegIII protein was detected specifically in epithelial cells at the crypt-villus junction, however after infection the number of PAP/RegIII-positive cells increased towards both the villus tip and the crypt bottom. So, a larger area of the surface epithelium was stained positive for PAP/RegIII after infection.
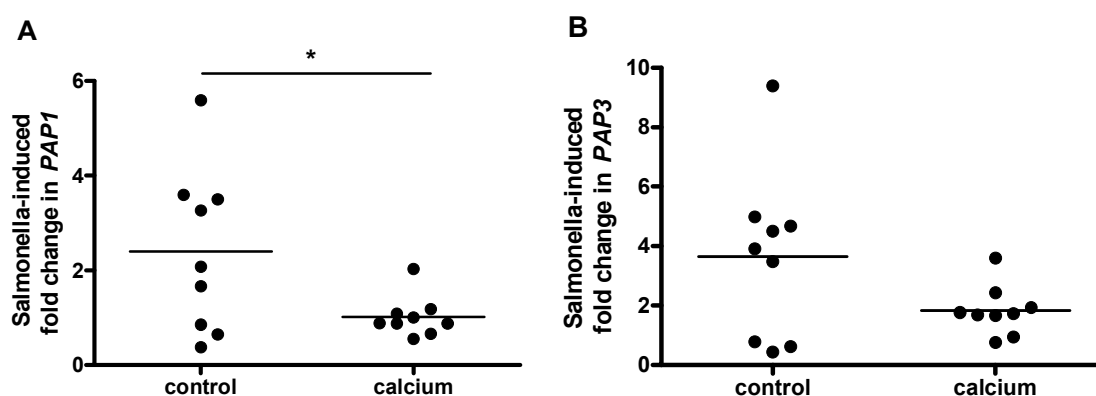
Figure 3. Relative pancreatitis associated protein 1 (*PAP1*, A) and pancreatitis associated protein 3 (*PAP3*, B) mRNA expression in ileal mucosa due to Salmonella infection of rats fed the control or calcium-supplemented diet. Individual values were first normalized to β-actin expression. These normalised data were used to calculate relative expression levels by setting the mean expression level of non-infected rats fed the identical diet at one. The mean of each diet group (n=9) is indicated by a line. The asterisk indicates p<0.05. Identical results were found when signals were normalised by comparison with aldolase mRNA (data not shown).

A

12.5x

B

40x

non-infected | **control** | **calcium**

infected

Figure 4. Pancreatitis associated protein/regenerating gene III (PAP/RegIII) protein expression in the ileal mucosa of non-infected and Salmonella infected rats fed either the control or calcium-supplemented diet. PAP/RegIII protein was present in epithelial cells at the crypt-villus junction (A, 12.5x; B, 40x). Immunostaining for lysozyme precisely identified paneth cells at the base of the crypts were PAP/RegIII staining was absent (data not shown). The number of PAP/RegIII positive cells increased after infection, especially in the direction of villus tips (A). Dietary calcium reduced PAP/RegIII expression in infected animals. No staining was observed in sections incubated without primary antibody (data not shown). All sections were counterstained with haematoxylin to visualise nuclei. (For full colour figure, see page 170).

Semi-quantification revealed that the percentage of total crypt-villus length stained positive for PAP/RegIII protein increased 34% after Salmonella infection in the control group ($p<0.05$; fig 5). In contrast, no significant increase in PAP/RegIII protein was observed in calcium-supplemented infected animals (fig 5). PAP/RegIII protein was not detected in the ileal Paneth cells, which was confirmed by specific staining of Paneth cells (data not shown).
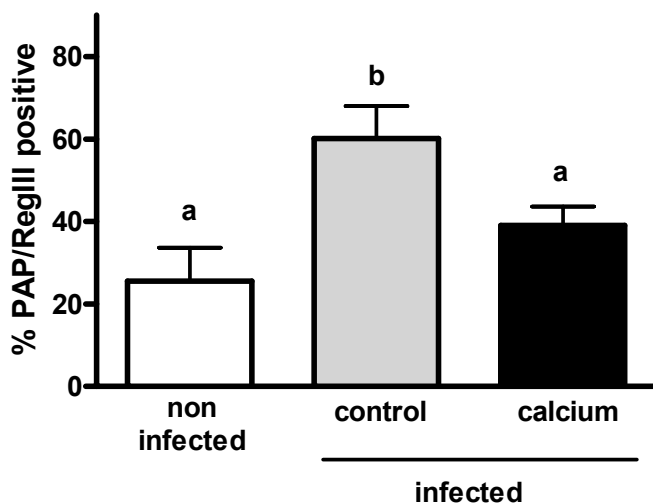
Figure 5. Pancreatitis associated protein/regenerating gene III (PAP/RegIII) protein expression in the ileum mucosa of non-infected (n=5) and Salmonella infected rat fed either the control (n=9) or calcium-supplemented (n=9) diet. Immunohistochemical slides (representative images are shown in fig 4) were used to quantify the percentage of total crypt-villus length stained positive for PAP/RegIII. Total crypt-villus length was not affected by infection or calcium. Results are expressed as means ± SE. Different letters indicate significant differences ($p<0.05$).

Presence and anti-microbial activity of PAP/RegIII in faeces

Immunoblotting revealed the presence of PAP/RegIII protein in faeces (fig 6A, lanes 1-10). Similar to the PAP/RegIII levels in the ileum mucosa, PAP/RegIII protein in faeces of rats fed the control diet was considerably increased from 3-4 days and remained steady until day 7 after infection. This infection-induced increase in PAP/RegIII was clearly less in calcium-supplemented infected animals. Although, the basal faecal PAP/RegIII level of the non-infected group was slightly higher in the calcium group in comparison with the control group.

Purified recPAP protein showed a single band of 15 kDa, which is identical to the size of PAP/RegIII detected in ileum mucosa (fig 6B, lanes 4 and 5). However, a second band of approximately 14 kDa was also apparent in the mucosal samples. When recPAP was added to faeces it was hydrolysed into two smaller forms. Adding recPAP to heat inactivated faeces largely prevented subsequent hydrolysation and merely preserved the 15 kDa band.

RecPAP in faecal water showed no bactericidal activity against the Gram-negative *S. enteritidis* or Gram-positive *L monocytogenes* (data not shown). Only growth of the latter was slightly inhibited (1 log CFU after 8 hrs incubation).
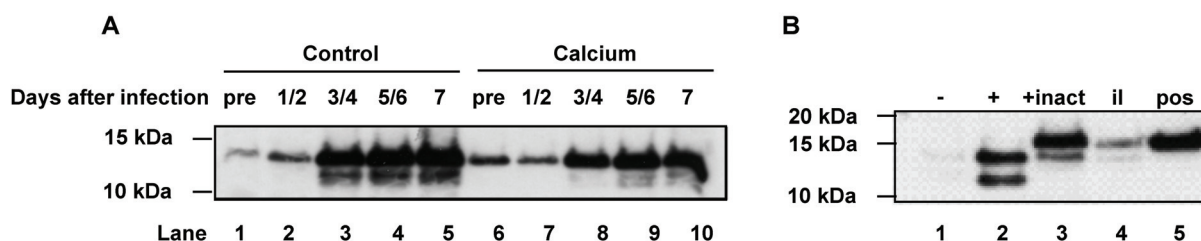
Figure 6. Effect of Salmonella infection and dietary calcium on pancreatitis associated protein/regenerating gene III (PAP/RegIII) excretion in faeces with time. Lyophilized faeces was pooled per group (n=9) per two days. SDSpage gels were loaded with equal quantities of faecal protein and analyzed for PAP/RegIII (A, lanes 1-10) by using specific PAP/RegIII anti-serum. Recombinant rat pancreatitis associated protein (recPAP) showed a single band (B, lane 5), similar to PAP/RegIII detected in the ileum mucosa of animals fed the control diet (B, lane 4). RecPAP incubated with pre-infection faeces from rats fed the control diet was hydrolysed into two smaller fragments (B, lane 2). Heat inactivation of faeces before addition of recPAP, largely prevented hydrolysis of the protein (B, lane 3). Non-spiked faeces (from non-infected rats) was added to lane 1 and served as negative control. No signal was detected when immunoblots were incubated without primary antibody (data not shown).

Discussion

In this study, we have shown that both PAP/RegIII mRNA and protein levels are upregulated in the ileal mucosa after oral Salmonella infection in rats. This upregulation was shown in parallel with infection severity as quantified by determination of intestinal Salmonella colonisation and translocation of this invasive pathogen. Furthermore, we demonstrated that the concentration PAP/RegIII protein present in faeces reflects abundance of this protein in the infected ileal mucosa.

In support of findings reported by Iovanna *et al*[7,18], our experiments clearly showed that PAP/RegIII protein was present in epithelial cells at the crypt-villus junction of the ileal mucosa. In contrast, others did report expression of *RegIII* in mouse ileal Paneth cells isolated by laser capture microdissecton before mRNA analysis.[8] In that study possible production of *RegIII* by

other mucosal cell types was not investigated or at least not described. Moreover, the latter study was performed in (ex-)germ-free animals, whereas, Iovanna's and our studies were performed in animals exposed to a conventional flora. When germ-free animals are exposed to micro-organisms for the first time this initiates a period of intensive contact between microbes and the gut mucosa, until their naïve immune system (for example secretion of sIgA) has evolved into a more mature one[29]. Therefore, results obtained in germ-free animals should be extrapolated with caution as those might not be relevant for the normal host with a conventional flora.

The increase of intestinal *PAP* or *Reg* mRNA expression after bacterial colonisation has not only been reported by our group[27]. Niewold *et al* identified *PAP1* mRNA upregulation by enterotoxigenic *E. coli*[21] and *Salmonella typhimurium* in a pig model.[22] These and our results indicate that *PAP/RegIII* expression is triggered by increased microbial-epithelial contact at mucosal surfaces. In order to establish infection, foodborne pathogens like *S. enteritidis* and enterotoxigenic *E. coli* bind to the small intestinal mucosa by using adhesins or colonisation factors which are known virulence factors[10,24,25]. Situations of increased host epithelium-microbial contact often coincide with gut mucosal inflammation, especially when pathogenic bacteria are involved[1]. Inflammation in general, is also suggested as causative agent of PAP/RegIII upregulation, as increased mRNA levels have been documented in inflamed colonic mucosa from patients with active inflammatory bowel disease[11,23] as well as in experimental models of colitis[23,32]. It should be realised that during active inflammation the mucosa might be damaged, its barrier function impaired, and mucosal cells are likely more exposed to micro-organisms from the gut lumen in comparison to healthy non-inflamed mucosa[20,30,31]. Hence, it is difficult to point out whether it is the inflammation or the microbe-host contact that induces PAP/RegIII upregulation. In a previous study we have shown that *PAP/RegIII* mRNA is not upregulated till day 3 after oral Salmonella infection[27]. As translocation of Salmonella has already occurred then[5,27], initial bacterial contact is less likely the main driver for PAP/RegIII upregulation. Here, we show that faecal PAP/RegIII, probably originating from mucosal cells, increased from day 3-4 after infection. Together these results suggest inflammation as trigger for PAP/RegIII increase. A study in which SPF animals are inoculated with an organism known to bind the gut mucosa without causing inflammation can show whether inflammation is indeed necessary for PAP/RegIII upregulation.

The function of PAP/RegIII is of most interest. Studies suggesting anti-microbial functions of PAP/RegIII have shown that mouse RegIIIγ was able to bind to Gram-positive bacteria *in vitro*[8]. Analogous to defensins, PAP/RegIII is a small protein (± 15 kDa) and has a secretion signal plus a carbohydrate-binding motif. Carbohydrates are present in the gut mucosa (for example mucines) and on the surface of bacteria. Up to now, actual excretion of PAP/RegIII to the gut lumen has only been slightly mentioned without showing data[18]. The results of our infection study clearly showed that PAP/RegIII is present in faeces and is upregulated in the same timeframe as the mucosal levels of this protein. This does not only suggest that faecal PAP/RegIII reflects ileal mucosal levels, but also suggests the gut lumen as functional site. In order to examine possible anti-microbial function, we investigated the bactericidal properties of PAP/RegIII as found by Cash *et al*,[8] however we studied this function in faecal water. The concentration of recPAP used in our experiments was 3.3 μM as this was approximately the concentration detected in faeces after Salmonella infection (estimated from immunoblots).

Moreover, this concentration fits well in the range applied in experiments of Cash *et al.*[8] We found that recPAP only slightly inhibited growth of Gram-positive *Listeria monocytogenes,* but absolutely no bactericidal activity was noticed against this pathogen, in contrast to studies described,[8] or against *S. enteritidis*. In the present study, we have shown that the protein was enzymatically hydrolysed in faeces with unknown effects on (anti-microbial) functionality. As hydrolysis is obviously relevant for the in vivo situation, we feel that our experiments using faecal water, including its enzyme activity, might be more relevant than experiments in clean buffers or media[8,14]. In view of the anti-microbial function, it is also remains unexplained why Gram-negative bacterial pathogens, like *S. enteritidis* (our present and previous[27,28] studies) and enterotoxigenic *E. coli*[21] are the most potent inducers of intestinal PAP/RegIII, whereas this mucosal protein has no anti-bacterial effect whatsoever against its inducers. Firstly, this might indicate that PAP/RegIII has this function at the mucosal interface only, where most of the protein could be intact. Secondly, it cannot be excluded that hydrolysis of PAP/RegIII in the intestinal lumen rather reflects activation than inactivation of a function yet to be discovered. Furthermore, if mucosal PAP/RegIII is secreted by epithelial cells it will face a myriad of carbohydrates associated with the mucosa. In that respect, binding of PAP/RegIII to mucins is an area to be investigated.

In conclusion, this study demonstrates that PAP/RegIII expression is significantly upregulated in rat ileum after oral Salmonella infection. Furthermore, PAP/RegIII protein levels present in faeces and the ileal mucosa both coincide with infection severity. As faecal PAP/RegIII levels can be monitored non-invasively and with time, its use as infection marker in animal and in human studies seems worthwhile to be further explored. Considering the growing interest in intestinal PAP/RegIII, elucidation of its functionality is scientifically very important and might be achieved by generation of PAP/RegIII knock-out animals. Very recently a PAP/HIP knock-out mice study on experimentally induced pancreatitis was published[13]. It would be worthwhile to address reaction of these mice to intestinal infection and inflammation as well.

Acknowledgements

References

1. Berkes, J., Viswanathan, V.K., Savkovic, S.D. & Hecht, G. Intestinal epithelial responses to enteric pathogens: effects on the tight junction barrier, ion transport, and inflammation. *Gut* 52, 439-451 (2003).
2. Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L., Van Doesburg, W., Witteman, B.J. & Van Der Meer, R. Diarrhea caused by enterotoxigenic Escherichia coli infection of humans is inhibited by dietary calcium. *Gastroenterology* 125, 469-476 (2003).
3. Bovee-Oudenhoven, I.M., ten Bruggencate, S.J., Lettink-Wissink, M.L. & van der Meer, R. Dietary fructo-oligosaccharides and lactulose inhibit intestinal colonisation but stimulate translocation of Salmonella in rats. *Gut* 52, 1572-1578 (2003).

4.   Bovee-Oudenhoven, I.M., Termont, D.S., Heidt, P.J. & Van der Meer, R. Increasing the intestinal resistance of rats to the invasive pathogen Salmonella enteritidis: additive effects of dietary lactulose and calcium. *Gut* 40, 497-504 (1997).

5.   Bovee-Oudenhoven, I.M., Termont, D.S., Weerkamp, A.H., Faassen-Peters, M.A. & Van der Meer, R. Dietary calcium inhibits the intestinal colonization and translocation of Salmonella in rats. *Gastroenterology* 113, 550-557 (1997).

6.   Bovee-Oudenhoven, I.M., Wissink, M.L., Wouters, J.T. & Van der Meer, R. Dietary calcium phosphate stimulates intestinal lactobacilli and decreases the severity of a Salmonella infection in rats. *J Nutr* 129, 607-612 (1999).

7.   Carroccio, A*., et al.* Pancreatitis-associated protein in patients with celiac disease: serum levels and immunocytochemical localization in small intestine. *Digestion* 58, 98-103 (1997).

8.   Cash, H.L., Whitham, C.V., Behrendt, C.L. & Hooper, L.V. Symbiotic bacteria direct expression of an intestinal bactericidal lectin. *Science* 313, 1126-1130 (2006).

9.   Christa, L*., et al.* HIP/PAP is an adhesive protein expressed in hepatocarcinoma, normal Paneth, and pancreatic cells. *Am J Physiol* 271, G993-1002 (1996).

10.   Cossart, P. & Sansonetti, P.J. Bacterial invasion: the paradigms of enteroinvasive pathogens. *Science* 304, 242-248 (2004).

11.   Dieckgraefe, B.K., Stenson, W.F., Korzenik, J.R., Swanson, P.E. & Harrington, C.A. Analysis of mucosal gene expression in inflammatory bowel disease by parallel oligonucleotide arrays. *Physiol Genomics* 4, 1-11 (2000).

12.   Folch-Puy, E., Granell, S., Dagorn, J.C., Iovanna, J.L. & Closa, D. Pancreatitis-associated protein I suppresses NF-kappa B activation through a JAK/STAT-mediated mechanism in epithelial cells. *J Immunol* 176, 3774-3779 (2006).

13.   Gironella, M*., et al.* Experimental acute pancreatitis in PAP/HIP knock-out mice. *Gut* 56, 1091-1097 (2007).

14.   Gironella, M*., et al.* Anti-inflammatory effects of pancreatitis associated protein in inflammatory bowel disease. *Gut* 54, 1244-1253 (2005).

15.   Graf, R*., et al.* Exocrine meets endocrine: pancreatic stone protein and regenerating protein--two sides of the same coin. *J Surg Res* 133, 113-120 (2006).

16.   He, W*., et al.* Bacterial colonization leads to the colonic secretion of RELMbeta/FIZZ2, a novel goblet cell-specific protein. *Gastroenterology* 125, 1388-1397 (2003).

17.   Iovanna, J., Orelle, B., Keim, V. & Dagorn, J.C. Messenger RNA sequence and expression of rat pancreatitis-associated protein, a lectin-related protein overexpressed during acute experimental pancreatitis. *J Biol Chem* 266, 24664-24669 (1991).

18.   Iovanna, J.L*., et al.* PAP, a pancreatic secretory protein induced during acute pancreatitis, is expressed in rat intestine. *Am J Physiol* 265, G611-618 (1993).

19.   Keim, V. & Loffler, H.G. Pancreatitis-associated protein in bile acid-induced pancreatitis of the rat. *Clin Physiol Biochem* 4, 136-142 (1986).

20.   Kleessen, B., Kroesen, A.J., Buhr, H.J. & Blaut, M. Mucosal and invading bacteria in patients with inflammatory bowel disease compared with controls. *Scand J Gastroenterol* 37, 1034-1041 (2002).

21.   Niewold, T.A., Kerstens, H.H., van der Meulen, J., Smits, M.A. & Hulst, M.M. Development of a porcine small intestinal cDNA micro-array: characterization and functional analysis of the response to enterotoxigenic E. coli. *Vet Immunol Immunopathol* 105, 317-329 (2005).

22.   Niewold, T.A*., et al.* The early transcriptional response of pig small intestinal mucosa to invasion by Salmonella enterica serovar typhimurium DT104. *Mol Immunol* 44, 1316-1322 (2007).

23.   Ogawa, H*., et al.* Increased expression of HIP/PAP and regenerating gene III in human inflammatory bowel disease and a murine bacterial reconstitution model. *Inflamm Bowel Dis* 9, 162-170 (2003).

24.   Pitman, R.S. & Blumberg, R.S. First line of defense: the role of the intestinal epithelium as an active component of the mucosal immune system. *J Gastroenterol* 35, 805-814 (2000).

25.   Pizarro-Cerda, J. & Cossart, P. Bacterial adhesion and entry into host cells. *Cell* 124, 715-727 (2006).

26.   Reeves, P.G., Nielsen, F.H. & Fahey, G.C., Jr. AIN-93 purified diets for laboratory rodents: final report of the American Institute of Nutrition ad hoc writing committee on the reformulation of the AIN-76A rodent diet. *J Nutr* 123, 1939-1951. (1993).

27.   Rodenburg, W., Bovee-Oudenhoven, I.M., Kramer, E., van der Meer, R. & Keijer, J. Gene expression response of the rat small intestine following oral Salmonella infection. *Physiol Genomics* 30, 123-133 (2007).

28.   Rodenburg, W*., et al.* Salmonella induces prominent gene expression in the rat colon. *BMC Microbiology* 7:84 (2007).

29.   Shi, H.N. & Walker, A. Bacterial colonization and the development of intestinal defences. *Can J Gastroenterol* 18, 493-500 (2004).

30.   Swidsinski, A*., et al.* Mucosal flora in inflammatory bowel disease. *Gastroenterology* 122, 44-54 (2002).

31.   Swidsinski, A*., et al.* Comparative study of the intestinal mucus barrier in normal and inflamed colon. *Gut* 56, 343-350 (2007).

32.   te Velde, A.A*., et al.* Comparative analysis of colonic gene expression of three experimental colitis models mimicking inflammatory bowel disease. *Inflamm Bowel Dis* 13, 325-330 (2007).

33.   Ten Bruggencate, S.J., Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L., Katan, M.B. & Van Der Meer, R. Dietary fructo-oligosaccharides and inulin decrease resistance of rats to Salmonella: protective role of calcium. *Gut* 53, 530-535 (2004).

34.   Ten Bruggencate, S.J., Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L. & Van der Meer, R. Dietary fructo-oligosaccharides dose-dependently increase translocation of Salmonella in rats. *J Nutr* 133, 2313-2318 (2003).

35.   Unno, M*., et al.* Structure, chromosomal localization, and expression of mouse reg genes, reg I and reg II. A novel type of reg gene, reg II, exists in the mouse genome. *J Biol Chem* 268, 15974-15982 (1993).

36.   Vasseur, S*., et al.* p8 improves pancreatic response to acute pancreatitis by enhancing the expression of the anti-inflammatory protein pancreatitis-associated protein I. *J Biol Chem* 279, 7199-7207 (2004).

37.   Watanabe, T., Yonekura, H., Terazono, K., Yamamoto, H. & Okamoto, H. Complete nucleotide sequence of human reg gene and its expression in normal and tumoral tissues. The reg protein, pancreatic stone protein, and pancreatic thread protein are one and the same product of the gene. *J Biol Chem* 265, 7432-7439 (1990).

38.   Zhang, H., Kandil, E., Lin, Y.Y., Levi, G. & Zenilman, M.E. Targeted inhibition of gene expression of pancreatitis-associated proteins exacerbates the severity of acute pancreatitis in rats. *Scand J Gastroenterol* 39, 870-881 (2004).

Chapter 5

A framework to identify physiological responses in microarray based gene expression studies: selection and interpretation of biologically relevant genes

Wendy Rodenburg[1,2,3 *]
A Geert Heidema[4,5,6*]
Jolanda MA Boer[4]
Ingeborg MJ Bovee-Oudenhoven[1,3]
Edith JM Feskens[6]
Edwin CM Mariman[5]
Jaap Keijer[1,2]

[1] TI Food and Nutrition, Wageningen, The Netherlands
[2] RIKILT-Institute of Food Safety, Wageningen, The Netherlands
[3] NIZO Food Research, Ede, The Netherlands
[4] National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands
[5] Department of Human Biology, Maastricht University, Maastricht, The Netherlands
[6] Division of Human Nutrition, Wageningen University, Wageningen, The Netherlands
* These authors contributed equally to this paper

Abstract

In whole genome microarray studies major gene expression changes are easily identified, but it is a challenge to capture small, but biologically important, changes. Pathway based programs can capture small effects, but may have the disadvantage to be restricted to functionally annotated genes. A structured approach towards the identification of major and small changes for interpretation of biological effects is needed.

We present a structured approach, a framework, that addresses different considerations in 1) the identification of informative genes in microarray datasets and 2) the interpretation of their biological relevance. The steps of this framework include gene ranking, gene selection, gene grouping and biological interpretation. Random forests (RF), which takes gene-gene interaction into account, is examined to rank and select genes.

For human, mouse and rat whole genome arrays, less than half of the probes on the array is annotated. Consequently, pathway analysis tools ignore half of the information present in the microarray dataset. The framework described takes all genes into account. RF is a useful tool to rank genes by taking interactions into account. Applying a permutation approach, we were able to define an objective threshold for gene selection. RF combined with Self-organizing maps identified genes with coordinated but small gene expression responses that were not fully annotated, but corresponded to the same biological process.

The presented approach provides a flexible framework for biological interpretation of microarray datasets. It includes all genes in the dataset, takes gene-gene interactions into account and provides an objective threshold for gene selection.

Introduction

Transcriptome analysis using whole genome microarrays is an elegant and widely used approach to identify the molecular mechanisms underlying dietary induced cellular or physiological changes. Both major effects as well as a wide overview of more subtle changes can be obtained. While the major differences are important for classification and identification of individual response genes, the smaller changes are an integral part of the physiological response and are essential for the identification of the physiological processes that are affected by the challenge or intervention. This is especially true in nutrition, where dietary interventions result in modest, but biologically important gene expression changes[1,12,32]. Also in the medical field it is increasingly recognized that the more subtle changes contribute importantly to outcome[29,30,37].

To translate microarray data into functional physiological information, a set of genes with the maximum amount of information and a minimum of noise is needed. Although a large number of methods exist to select genes from microarray datasets, most methods aim to identify the smallest possible set of genes that still can discriminate, for example to classify malignancies, predict therapeutic outcomes or diagnose physiological responses[7,37]. These methods may not always be appropriate to select larger set of genes for biological interpretation, that includes the smaller changes. These smaller changes are part of the response to medication or disease, which occurs through the interactions of multiple genes, via signaling pathways or other functional relationships. Small changes, variability among individuals and the often small samples sizes on the one hand and the large number of genes tested on the other, make it difficult to distinguish true differences from noise[29,48]. Careful planning and execution of microarray experiments nowadays offers technically high quality data, with a minimum of noise. However, the combination of small gene expression changes and the needed selection of a largest informative set of genes demands sophisticated selection methods. A structured framework that incorporates the different considerations in the identification of informative genes and the interpretation of their biological relevance is needed. Here, we describe the steps of such a framework and address the following considerations: gene ranking, gene selection, gene grouping and biological interpretation.

Gene ranking

To identify genes of relevance within the total dataset, genes are ranked by a measure of importance. As such, fold change has often been used. However, fold change is not a reliable measure, as this measure does not take variability in the data into account[2,46]. Therefore, other measures that do take variability into consideration should be used. The most commonly used approach for gene selection in two-class microarray studies that takes variability into account is the conventional t-test, while ANOVA is used for multi-class studies. Genes are tested independently and a p-value is assigned to each gene, which can be used to rank genes by their importance. However, by ranking genes by a univariate test-statistic such as the t-statistic, all genes in the dataset are assumed to be independent and gene-gene interactions are not taken into account. In biological responses, gene-gene interactions will take place as these responses often result from co-regulation of genes[4,39]. Consequently, by testing each gene independently, weak to

small genetic effects that only in interaction make an important distinction between different study groups will not be detected by using a univariate test.

Gene selection

For functional interpretation the total ranked gene set can be used, but this will include noise and selection of the most important genes is needed. The difficulty in gene selection is how to define the threshold. The threshold to select the differentially expressed genes influences the functional interpretation. Selection of genes is to some extent subjective, as there are no clear thresholds for existing methods. For t-test, the threshold choice is flexible and the significance level is chosen by the researcher[3,8]. However, a threshold should preferably be defined in an objective way. Procedures can be applied to correct for multiple testing, such as the family-wise error rate (FWER) or the false discovery rate (FDR)[19,44]. However, these procedures can be overly stringent, resulting in identification of only the most important changes and possibly discarding other relevant genes[30].

Gene Grouping

Each probe on a microarray corresponds to a specific nucleotide sequence, which represents a specific gene. Most genes known to be involved in a functional category are annotated in annotation databases, such as the GO-database[16], KEGG[21] or Entrez Gene[22]. Whole genome microarrays contain annotated genes as well as non-annotated genes. Although the extent to which spots on whole genome microarrays are annotated has not exactly been established, many known genes are not annotated in functional analysis tools, for example GO-annotated, and are thus lost for biological interpretation when a pathway program uses the GO-database as source [14,22]. However, the non-annotated genes may provide important new targets. Clues on the function of these genes can be obtained by establishing similarities in expression behavior to known genes. Genes with similar gene expression can be identified using self-organizing maps (SOM) and hierarchical clustering[34,43,45]. Grouping based on similarity in expression behavior is also useful for functional interpretation of known genes.

Biological interpretation

Biological interpretation is the final step in this framework. A useful way to interpret microarray data is pathway analysis. In pathway analysis the effect of treatment on biological processes or co-regulated gene sets are studied, rather than effects on individual genes[22,47]. A commonly often used approach is to import a list of genes, that meets the threshold criteria, into a pathway program, such as freely available ErmineJ, GeneMapp, David/EASE or commercially available ones like Metacore or Ingenuity. These programs search through public or private databases to link related genes that are grouped in biological processes.

Recently, new methods have been developed for functional interpretation which circumvent the need to preselect genes[36]. One of these methods is Gene Set Enrichment Analysis (GSEA)[42]. This method enables detection of important pathways where all genes in a predefined set (for instance a GO-category) change in a coordinated manner[28,29]. This is highly relevant for studies where subtle, but coordinated changes in expression can be expected. However, GSEA may have the disadvantage that it is restricted to, and therefore only informs about, functionally annotated

genes. Thus, not all information that is available in the dataset is used. Nevertheless, the application of GSEA has shown that small effects can be captured when coordinate gene expression changes are taken into account[29].

In this study we describe a framework for functional interpretation of microarray based expression studies using two real gene expression datasets. For gene ranking and selection, we have examined the usefulness of Random forests (RF)[6]. RF is one of the statistical methods that have been developed to select genes from large datasets containing many variables in small sample sizes. RF and other supervised methods like Support Vector Machines (SVM) and Discriminant Analyis (DA) have mainly been used to select genes that provide the best classification performance for diagnostic purposes (e.g.[20,38]). In a simulation study, RF performed better than an univariate test[26]. In microarray studies, RF was shown to outperform other classification methods, especially when the number of classes is moderate[13,24]. RF could also be a suitable tool to rank and select a larger subset of genes for further interpretation, as it is has many advantages[13]. One major advantage of RF is that it provides an importance measure for each gene, which can be used to rank the genes. Furthermore, the advantage of this importance measure is that it takes gene-gene interactions in the ranking of genes into account. In this way, RF is able to capture not only the main effects in a dataset, but also the variables with weak to small genetic effects that mainly contribute by interactions with other genes. Interaction between genes increases the importance of the individual interacting genes, making them more likely to be given high importance relative to other genes. Genes with a higher importance index are more associated with differences resulting from the treatment. As RF takes gene-gene interactions into account in the ranking of genes, this method was applied within this framework as a tool to rank genes at the first step. However, RF does not provide a threshold to define which genes should be selected for further interpretation. Therefore, after applying RF to rank genes by their importance index, we examined an approach to define a threshold for the genes ranked by RF to select biologically important genes in an objective way. After selection, genes were clustered by self-organizing maps (SOM) which clusters genes with similar gene expression in ordered profile-groups. The advantage of combining results obtained with SOM and information obtained at previous steps is that insight can be obtained whether genes within the same profile contribute by their main effect and/or that interaction effects are present. Finally, for each gene expression dataset, the selected genes obtained by RF were incorporated in pathway programs (Metacore and ErmineJ) and compared to the results obtained with GSEA. Together this provides a stepwise framework focusing on the different considerations in the identification of informative genes and the interpretation of their biological relevance.

Methods

Datasets

To illustrate and examine the framework considerations, we have used two whole genome gene expression datasets obtained from the same dietary study. In this study, two groups of Wistar rats were fed different diets for 2 weeks. One group of rats received a control diet (n=12) and the

other an experimental diet (n=12). The experimental diet is identical to the control diet, but additionally contains fructo-oligo saccharides. Detailed analysis of the effects of the diet is subject of another paper. The two datasets are obtained from two different tissues, colon and cecum. RNA from colon mucosa and cecum mucosa was isolated, reverse transcribed into cDNA, labeled and individually hybridized to Agilent-Whole Rat Genome Microarrays (G4131A). Labeling was performed by incorporating Cy5 for individual samples and Cy3 for pooled RNA. Hybridization and washing were carried out according to Agilent protocols. A total of 24 arrays for colon were analyzed, one array did not pass the quality controls based on MA plot and signal intensity distribution [2,40]. Therefore, the colon dataset contained 23 arrays in total. The cecum dataset contained 22 arrays in total, since two cecum RNA samples were excluded based on poor quality of RNA. We preprocessed the microarray datasets as described [33]. Only genes with an average signal 1.5 times above the background were taken into account for further data analysis, equal to 28180 genes for colon and 21049 genes for cecum. Gene expression values were log-transformed before performing statistical analyses. The data have been deposited in NCBIs Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) and are accessible through GEO Series accession number GSE5943.

Statistical analyses
T-test

T-tests to obtain t-statistics and corresponding p-values for the differences in mean gene expression between the two treatment groups were performed using the program GeneMaths XT (Applied Math, Sint-Martens-Latem, Belgium). Within the same program FDR analyses according to the Benjamini and Hochberg procedure [19] were performed.

Random forests

In RF a group of tree-based models (the forest) can be used to rank genes with an important contribution to the treatment variable. Each tree starts with the total dataset, which is recursively splitted into smaller and more homogeneous groups to fit models for predicting the different treatment groups from the selected genes. Within the forest, different trees are obtained by bootstrapping and random subset selection. In more detail, each tree is constructed from a bootstrap sample of the total dataset. A bootstrap sample is a sample of observations (for example rats) from the original dataset with replacement. The bootstrap sample therefore contains the same number of observations as the original dataset, but some observations are sampled more than once, while others are left out. The sampled observations are used to construct the tree, whereas each left-out sample is used to obtain a prediction of that tree to what extent the left-out sample can be appointed to one of the groups based on their genes. A prediction for the forest is obtained by aggregating the predictions over all trees for which the sample was left out. The prediction error of the forest is then the proportion of misclassified samples and is a quantitative measure of the quality of the prediction. For each split in a tree, the gene that gives the best split is not selected from the total set of genes, but from a random subset of genes. The number of randomly selected genes that is used to be searched through for the best split is referred to as m_{try}. RF performance is usually not sensitive to this parameter and it is suggested to use $\sqrt{total\ nr\ of\ genes}$ as a default value for m_{try} [6,25]. Comparing the default value and

values lower and higher than the default for both colon and cecum, we obtained similar prediction errors for different m_{try} values (data not shown). Therefore, default values for m_{try} ($\sqrt{total\ nr\ of\ genes}$) were chosen for both colon (167 genes) and cecum (145 genes) to perform the RF analyses.

More important genes will discriminate better between the treatment groups and will therefore be present in most of the trees and more often selected at a split close to the total sample. On the other hand, less important genes will be less present in the different trees and selected at splits farther from the total sample. Importance of genes is defined by a measure referred to as the importance index (I_m). This Im is obtained by comparing the predictive performance of the forest for all genes with the predictive performance of the forest in which the values of one gene are randomly permuted in the trees for the left-out samples. For more important genes this will results in larger differences in the predictive performance as these genes are selected at splits close to the total sample while less important genes that are selected at splits farther from the total sample. The more important genes will therefore be given a larger importance index. By permuting the values for one gene, not only the effect of this gene is taken into account, but also all possible interactions of this gene with other genes. Interactions between genes increase the importance index for each of the genes that are part of the interaction. In this way random forests takes interactions between genes into account. Genes are ranked according to their importance. To obtain stable estimates of the I_m, large numbers of trees in the forest are needed [25,26]. Since one tree will only be able to capture the interactions between the genes present in that tree the use of large numbers of trees allows capture of many interactions. Therefore we set the number of trees grown in the forest to 40000. We used all genes in the dataset in the analysis and the I_m values for each gene was used as measure to rank the genes.

To obtain a threshold for selection of genes for subsequent interpretation, the permutation test [9,27] was applied. We used 100 permutation datasets, in which the group labels are randomly permuted. For each permutation dataset, RF analysis was performed using the same parameter settings as for the observed dataset. Next, for each permutation dataset I_m's for the genes were obtained and genes were ranked. The distribution of the I_m values derived from the permutation datasets indicates how the I_m's of the genes behave in the absence of a true association with the treatment. To define the threshold for selecting genes, two approaches were taken. The first approach was to determine the value of I_m where the I_m of the observed dataset was equal to, or lower than the I_m for at least one of the 100 permutation datasets. This corresponds to a significance level of $p < 0.01$. The second approach to define the threshold (explained and illustrated at the GeneSrF website[17]) was to determine the number of genes with I_m larger than the mean value of I_m for the first ranked gene obtained from the 100 permutation datasets. However, this second approach yielded a very small number of genes, 11 for colon and 19 for cecum, which was very unlikely compared to the results of other methods (corresponding p-value for these threshold were $p = 7 \cdot 10^{-7}$ for colon and $9 \cdot 10^{-6}$ for cecum). Therefore, we only used the results of the first approach.

To examine whether RF provides reproducible results over different analyses, we performed several analyses (runs), each time using the same parameter settings, but a different seed value. The seed value controls the random number generator and different seed values generate different forests. The results can be repeated if the same seed value is used. We examined the

reproducibility of RF by comparing the I_m of the genes for different runs. Each run can return slightly different results as in RF each tree is constructed on a bootstrap sample of the observations (rats), and at each split of the tree the best discriminating gene is selected from a random subset of genes (m_{try}).

The permutation test that was used to determine the threshold of the I_m was also used to obtain the significance of the prediction error of the random forest model. For each permuted dataset, a prediction error was obtained by random forests. The proportion of permutation datasets with a prediction error equal to or lower than the prediction error of the random forest model of the observed dataset provided the significance of the model.

Software for Random Forests (RF) is freely available, including R-packages[10,25,35,41], and the original Fortran code[5]. For analyses with RF we have applied the R-package RandomForest to obtain the I_m for the different genes.

Gene grouping: SOM

For the gene sets selected with the obtained RF threshold (935 genes in colon, 165 genes in cecum), self-organizing maps (SOM) analyses were performed. SOM was chosen because this method clusters genes with similar expression into ordered groups, rather than random groups, of genes with similar expression profiles. The number of SOM-profiles was set to 90 for colon and 16 for cecum, corresponding to an average of approximately 10 genes per profile-group. To distinguish between genes that mainly contribute by their interaction effect or their main effect, genes selected by RF were compared to the same number of genes ranked by t-test. We explored whether profiles consisting of genes only selected by RF were present, which indicate profiles consisting of gene-gene interaction effects.

To perform SOM-analysis, both commercial (e.g. GeneMaths XT) and free open-source software packages (e.g. Orange machine learning software[11] at http://www.ailab.si/orange) are available. In this study we used GeneMaths XT (Applied Math, Sint-Martens-Latem, Belgium) to obtain the SOM-profiles.

Biological interpretation: Pathway analysis

For the genes selected by RF, we performed pathway analyses for biological interpretation. The pathway results obtained for genes selected by RF were compared with pathway results obtained for the same number of genes selected by t-test, to assure comparability. For pathway analysis we used the freely available software ErmineJ [23] and the commercial program Metacore[15]. ErmineJ is a web-based application for identification of Gene Ontology (GO) processes on input gene sets. Metacore is a package of GeneGo (GeneGo Inc, St. Joseph, MI).

In ErmineJ we used over-representation-analysis (ORA), in Metacore GO-processes were used for pathway analysis. For both ErmineJ-ORA analysis and Metacore-GO processes, gene sets existing of 5-250 genes were tested. In both analyses, gene lists selected by RF or t-test were classified into GO processes. These processes were ranked according to their p-value, which represents the probability that a particular process is selected by chance. Each pathway program uses different statistical tests to calculate these probabilities, this is beyond the scope of this paper and is discussed in Goeman *et al*[18]. For both programs we selected pathways with two selection

criteria i) the pathways should have a p<0.001 and ii) the pathways should include at least three selected genes.

We also analyzed which biological pathways were enriched using GSEA[42]. In GSEA, enrichment of genes in a gene set is based on the ranking of the genes within the whole dataset[36]. We included functional c2 gene sets originated from KEGG, GenMapp and BioCarta with 5-500 genes with FDR q-value <0.25 and ranked on NES score and nominal p-value.

Results

Whole genome arrays are not fully examined in pathway analysis programs

Whole genome microarray analysis combined with pathway analysis is an attractive approach to identify effects of an intervention, but the analysis is limited to those genes that are annotated in database used by the program. To assess completeness of annotation, we examined first the extent to which genes were incorporated in the analysis in three different pathway programs, Metacore (GeneGo), ErmineJ and GSEA. This was performed for the two most widely used array platforms, Agilent and Affymetrix, and for three different species: human, mouse and rat. Only 23-48% of the probes on whole genome microarrays are translated to functional categories by these programs (Table 1). Ermine J is not included as it does not provide the number of incorporated genes. Annotation in this program is based on the specific GO-term(s) linked to the gene, which for the Agilent 44k Rat array applies to 7437 genes (18%). Altogether, analysis only based on functional annotation and co-occurrence in genesets, leaves out at least half of the microarray data, and thereby potential new targets.

Table 1. Percentage of probes from whole genome microarrays identified by the pathway programs Metacore and GSEA.

Number of probes imported		Number of probes linked to program database			
		Metacore*		GSEA[†]	
		Number	Percentage	Number	Percentage
Agilent					
Human	41675	12976	31	17517	42
Mouse	41534	13714	33	19589	47
Rat	41372	9489	23	14631	35
Affymetrix					
Human	54675	22792	42	20606	38
Mouse	45102	18105	40	21891	48
Rat	44761	12259	39	13342	43

* Spots linked to a GO-term

† Spots linked to a gene symbol

Information content of gene expression datasets

In both gene expression datasets (p=28180 for colon, p=25049 for cecum) the extent of differential gene expression induced by the dietary treatment was small; in colon, 179 genes were differentially expressed with a change of more than 1.5 fold, while in cecum the number of differentially expressed genes was 164. Based on fold change the datasets are similar in number of expressed genes and magnitude of differential expression (fold change). However, the two datasets differed in the significance of expression, with the colon dataset containing substantially

more significantly differentially expressed genes (table 2). With a t-test threshold of p<0.001, 803 genes were differentially expressed in the colon dataset, while 123 genes were differentially expressed in cecum. Application of FDR using a threshold of q<0.01 resulted in selection of 231 genes in colon and 19 genes in cecum. RF models were found to be significant in both colon (p<0.02) and cecum (p<0.01), indicating that gene expression differences were present.

Table 2. Characteristics of the colon and cecum dataset

	Total number of genes in dataset	Fold change >1.5 fold*	T-test p<0.001	FDR q<0.01
Colon	28180	179	803	231
Cecum	25049	164	123	19

*Fold change experimental diet/control diet.

Gene ranking: taking gene-gene interactions into account

Genes were ranked according to their I_m obtained using RF. To obtain insight in the ranking of genes by RF, we compared the results from RF with the ranking of genes by the commonly used t-test. For the genes present in the dataset the absolute values for the t-statistics are plotted against the I_m of RF (see Figure 1). In both datasets, I_m obtained from RF does show a similar trend with t-statistics. Both RF and t-test rank genes in common (Figure 1, box A), indicating strong gene effects related to the treatment. Genes ranked high by RF, compared to t-test (Figure 1, box B), are indicative of weak gene effects that are likely to be related to the treatment in interaction with other genes.
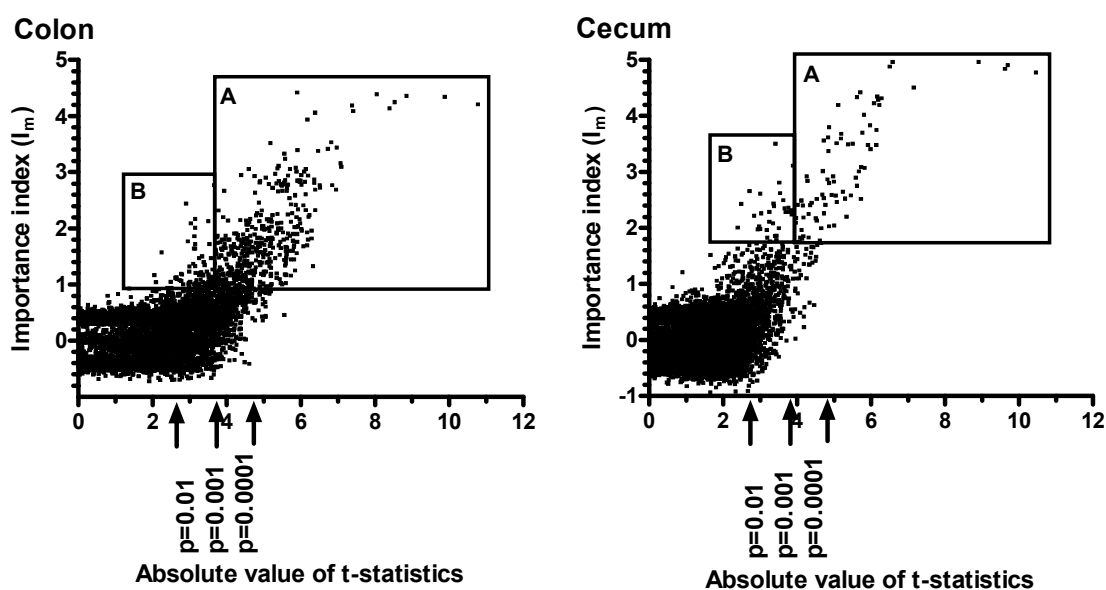
Figure 1. Plot of absolute value of t-statistics against I_m for colon (left) and cecum (right) dataset. Box A: Genes highly ranked by both RF and t-test. Box B: Genes highly ranked exclusively by RF.

Gene selection: defining an objective threshold

We aimed to define an objective threshold for I_m by using a permutation approach (see method section). This permutation test provides an indication where noise starts to interfere with real gene effects. For both colon and cecum the highest ranked genes from the observed dataset had higher I_m values than the ranked I_m values obtained from the permuted datasets (see Figure 2). To define the threshold, we determined the I_m value where genes in the observed dataset have equal

or higher I_m values relative to the genes in the permuted datasets. The point where the I_m values of the observed dataset equaled that of at least one of the 100 permutated datasets was chosen as threshold, which is equal to a significance level for the I_m of $p<0.01$.
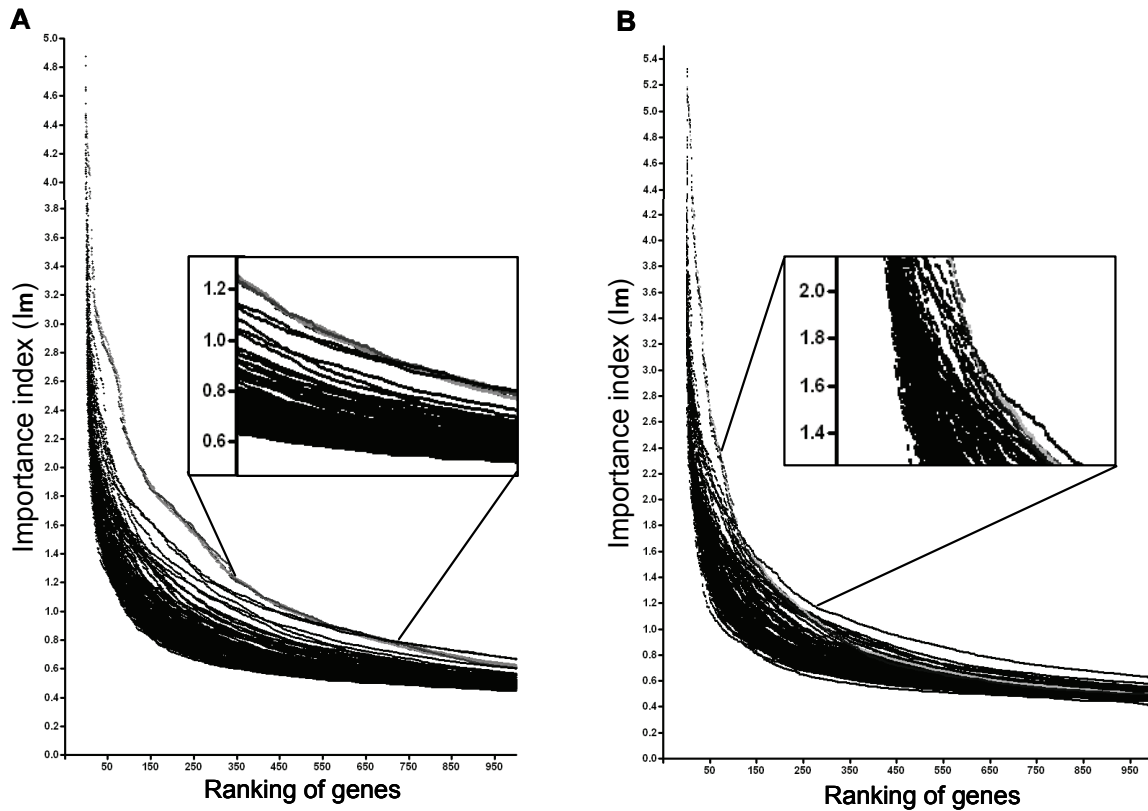
Figure 2 a and **b**. Genes, of 100 random sets (black lines) and real sets with different seed values (colored lines), ranked by the I_m values. For colon (A) and cecum (B) datasets. (For full color figure, see page 171).

We performed 15 runs (each with a different seed value) resulting in very similar thresholds (results not shown). For colon a mean threshold of $I_m=0.906$ and for cecum a mean threshold of $I_m= 1.753$ was obtained. For each run, the genes with I_m values above the threshold were determined. Genes with higher I_m values were always selected over the different runs. However, genes with ranking close to the threshold (lower I_m values) were not selected over all runs, thus the selection of these genes varied between different runs. We chose to include all genes that were selected in at least one run, and not only the overlapping genes, because the number of genes that were additionally selected over increasing numbers of runs decreased rapidly (figures 3a and b, tables 3a and 3b, for colon and cecum, respectively). This likely indicates that additionally selected genes are truly affected by the treatment and not randomly selected noise. After 10 runs for colon and 11 runs for cecum, the proportion of genes additionally selected became and remained less than 2%. Therefore, more runs were not performed. Combining the results of different runs resulted in the selection of 935 genes above the threshold for colon and 165 genes above the threshold for cecum. These genes were selected as the set of genes being related to the treatment.

A

B
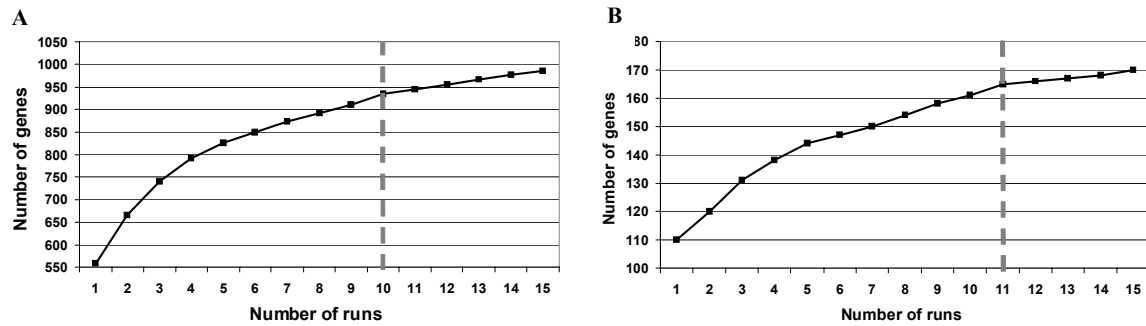
Figure 3. Genes selected by RF thresholds Im> 0.906 for colon and Im>1.753 for cecum. The total number of selected genes is plotted against the number of runs.

Table 3. Selection of genes by RF threshold.

A: Colon

Run	Number of genes*	Total number of genes†	Genes added	
			Number	Percentage
1	558	558	-	-
2	552	665	107	19.2
3	558	740	75	11.3
4	558	791	51	6.9
5	557	825	34	4.3
6	562	849	24	2.9
7	542	873	24	2.8
8	557	891	18	2.1
9	564	911	20	2.2
10	549	**935**	24	2.6
11	560	945	10	1.1
12	554	955	10	1.1
13	540	966	11	1.2
14	573	977	11	1.1
15	547	985	8	0.8

* Number of genes selected with threshold I_m>0.906.

† The number of genes selected after each additional run.

B: Cecum

Run	Number of genes*	Total number of genes†	Genes added	
			Number	Percentage
1	110	110	-	-
2	109	120	10	9.1
3	118	131	11	9.2
4	112	138	7	5.3
5	111	144	6	4.3
6	108	147	3	2.1
7	112	150	3	2.0
8	108	154	4	2.7
9	112	158	4	2.6
10	115	161	3	1.9
11	111	**165**	4	2.5
12	114	166	1	0.6
13	108	167	1	0.6
14	115	168	1	0.6
15	108	170	2	1.2

* Number of genes selected with threshold I_m>1.753.

† The number of genes selected after each additional run.

Comparison of gene selection by RF, t-test and fold change

Genes selected based on the RF threshold (935 genes in colon and 165 genes in cecum) were compared with an equal number of genes selected by t-test. For t-test this resulted in inclusion of genes with p<0.0014 (q<0.04) for colon and p<0.0018 (q<0.23) for cecum. In colon 679 genes (72.6%) and in cecum 112 genes (67.9%) overlapped between RF and t-test. As shown in the volcano plots (figures 4a and b), gene sets selected by RF include the most significant genes based on t-test, as was also seen in figure 1. Furthermore, the volcano plots show that RF and t-test also differ in selection of genes. Several genes with high fold change, that would not have been selected based on t-test alone, are also selected by RF.

For both datasets, the set of selected genes by RF were used for subsequent gene grouping and biological interpretation.
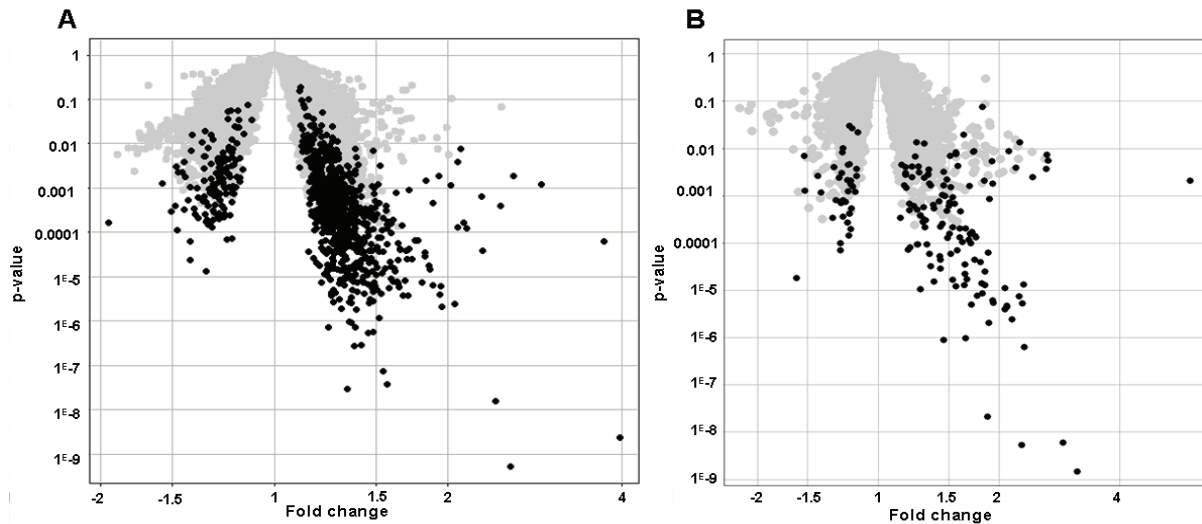
Figure 4. Volcano plots for colon (A) and cecum (B). Fold change is plotted against p-value. All genes are shown, genes selected by RF are shown in black (935 for colon, 165 for cecum).

Gene grouping: obtaining gene expression profiles by SOM

For grouping of the genes selected by RF, we applied self-organizing maps (SOM)-analysis, to find groups of highly correlated genes. While SOM is mostly used to identify patterns in time or as a result of multiple treatments [43], it will also identify patterns of coordinate changes over a number of animals. In figure 5, the groups of genes with similar expression are shown for colon and cecum respectively. For both colon and cecum, profiles are present that consist mainly of genes that are selected exclusively by RF (light gray). SOM analyses for genes selected by the t-test did not result in profiles consisting of genes exclusively selected by t-test (data not shown). Apparently, RF selects genes with main effects similarly to the t-test, but additionally selects genes (not selected by t-test) that can be grouped in profiles, which are likely to be related to the treatment by gene-gene interaction effects.
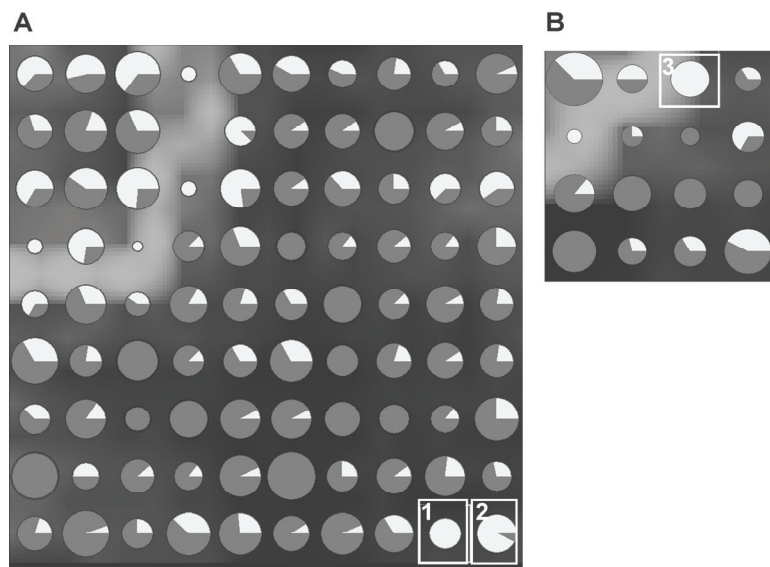
Figure 5. SOM profiles for colon (a) and cecum (b). The total number of SOM-profiles was arbitrarily set to 90 for colon and 16 for cecum, corresponding to an average of approximately 10 genes per profile. The size of the circles corresponds to the number of genes included in the group (range of genes per profile: colon 1-19, cecum 2-27). Within each profile, genes that overlap between RF and t-test are shown in dark gray, and genes exclusively selected by RF are shown in light gray. Genes in profiles 1, 2 and 3 were analyzed in more detail.

We examined whether the genes exclusively selected by RF and highly enriched within one profile shared similar biological functions. Herefore we selected profiles consisting of mainly RF selected genes. For colon two profiles and for cecum one profile was selected (figure 5 white squares). The first colon profile (profile number 1) consisted of nine genes, four genes with unknown function (ESTs) and five genes which were annotated but not classified to a known GO-process. After literature and database search these five genes could not be linked to a single biological process (table 4a). The second colon profile (number 2) consisted of 13 genes, of which 12 were only selected by RF. Five genes were annotated in a GO-process (bold gene names in table 4b), of which four are part of the same GO-process: cellular component organization and biogenesis. The remaining eight genes consisted of two ESTs, and six genes which are presently poorly understood, as further database and literature mining did not reveal a relation to a known biological process. One of these 6 (palladin) was recognized to play a role in maintaining normal actin cytoskeleton architecture [31], indicating a possible role in the same biological process as the 4 annotated genes within this SOM profile.

The cecum profile consisting of exclusively RF selected genes (profile number 3) consisted of 13 genes, comprising ten unique genes. Three of the ten genes were annotated by GO, of which two are part of the GO-process immune response. Further database and literature mining revealed that six of the seven other genes had a function related to immune response (table 4c). This confirms the notion that genes with a similar expression profile selected from a micorarray dataset exclusively by RF may be enriched in the same biological process. It further indicates that this is a strategy to hunt for biological function of genes and to reveal new biological processes related to treatment

Table 4. Genes mainly selected exclusively by RF, grouped in SOM profiles (white boxes figure 5).

A: Colon SOM profile number 1.

Gene name*	Sequence ID	Gene symbol[†]	FC[‡]	p-value	q-value
3222401M22Rik protein (LOC363231)	XM_343571	-	1.39	0.005	0.08
2410024A21Rik protein (LOC314415)	XM_234506	-	1.21	0.014	0.13
Rattus norvegicus cDNA clone UI-R-A1-dv-f-02-0-UI 5'	BF558849	-	1.20	0.007	0.10
Uronyl-2-sulfotransferase	XM_341728	*Ust*	1.12	0.074	0.31
Ring finger protein 10	XM_213797	*Rnf10*	1.21	0.026	0.18
Midnolin	TC480469	*Midn*	1.31	0.002	0.05
Mitsugumin 29	XM_342316	*Mg29*	1.16	0.012	0.13
Carbonic anhydrase I (Carbonate dehydratase I)	XM_226922	*Ca1*	1.28	0.007	0.10
Polyglutamine-containing protein	BF546374	-	1.22	0.001	0.04

* None of these genes were annotated by GO.

† Genes without official gene symbol are indicated with -.

‡ Fold change experimental diet/control diet.

B: Colon SOM profile number 2.

Gene name*	Sequence ID	Gene symbol[†]	FC[‡]	p-value	q-value
Hypothetical protein FLJ32871	XM_219819	-	1.26	0.025	0.18
GCD14/PCMT domain containing protein	NM_001007706	-	1.23	0.015	0.14
Telomeric repeat binding factor 2	XM_341683	*Terf2*	1.26	0.002	0.05
Probable nocturnin protein	XM_344988	-	1.11	0.184	0.48
cGMP-dependent protein kinase 1, beta isozyme	XM_219807	*Prkg1*	1.21	0.010	0.11
High mobility group nucleosomal binding domain 1	BI303604	*Hmgn1*	1.11	0.156	0.44
Cyclin-dependent kinase 5	NM_080885	*Cdk5*	1.18	0.021	0.16

Beta-sarcoglycan	XM_223355	*Sgcb*	1.11	0.096	0.35
Phosphodiesterase isoform	AF053097	*Pde*	1.18	0.005	0.08
Palladin, cytoskeletal associated protein	XM_214338	*Palld*	1.20	0.010	0.11
splA/ryanodine receptor domain and SOCS box containing 3	XM_220230	*Spsb3*	1.20	0.003	0.07
Kinesin family member 5B	XM_341538	*Kif5b*	1.24	0.001	0.03
Acyl Transferase	XM_235527	*Mct*	1.13	0.010	0.11

* Genes annotated by GO are presented in bold.

† Genes without official gene symbol are indicated with -.

‡ Fold change experimental diet/control diet.

C: Cecum SOM profile number 3.

Gene name*	Sequence ID	Gene symbol†	FC‡	p-value	q-value
Anti-NGF30 antibody light-chain , variable and constant regions	U39609	-	2.65	0.005	0.31
Ig germline kappa-chain gene C-region	M12981	*Igkc*	2.63	0.007	0.34
Immunoglobulin joining chain	XM_341195	*Igj*	2.11	0.009	0.35
Immunoglobulin rearranged κ-chain mRNA variable (V) region	CO562777	*Igkv*	1.92	0.005	0.31
Anti-acetylcholine receptor antibody gene, κ-chain, VJC region	L22655	-	2.25	0.013	0.39
Ig germline kappa light chain joining (J) segments	J00746	*Igkjca*	1.72	0.009	0.35
Periostin, osteoblast specific factor	XM_342245	*Postn*	1.93	0.002	0.23
Immunoglobulin kappa light chain variable region	AF217591	*Igkv*	1.73	0.009	0.35
Chemokine (C-X-C motif) ligand 12	NM_022177	*Cxcl12*	1.63	0.020	0.46
Ig active kappa-chain mRNA VJ-region from immunocytoma IR162	M15402	*Igkac*	2.20	0.004	0.28
		Igkac	1.56	0.008	0.35
		Igkac	1.56	0.008	0.34
		Igkac	1.58	0.004	0.29

* Genes annotated by GO are presented in bold.

† Genes without official gene symbol are indicated with -.

‡ Fold change experimental diet/control diet.

Biological interpretation: Pathway analysis to obtain biological processes

To examine whether pathway programs are able to identify differences between RF selected genes and t-test selected genes, we applied pathway analysis for the set of genes selected by RF and compared this with the same number of genes selected by t-test (935 genes for the colon dataset, and 165 for the cecum dataset). To ensure that we covered different pathway analysis methods, we used two pathway programs, Metacore and ErmineJ. For both colon and cecum the comparison between RF and t-test based selection showed highly comparable results per pathway program (Metacore table 5a and 6a, EmineJ table 5b and 6b). However, the ranking of processes was somewhat different and each selection method (RF or t-test) identified some unique processes.

GSEA does not require preselection of genes, although information may be lost due to incomplete annotation. GSEA is especially suited to identify processes based on interaction. To see whether similar or complementary information is obtained, we analysed the complete colon and cecum datasets with GSEA. We focused on pathway related GSEA genesets, obtained from GO, GenMapp and Biocarta, to allow for comparison. Only few genesets were found to be significantly enriched (FDR<0.25 according to GSEA), 12 in colon and 6 in cecum. The small number of processes identified by GSEA analysis suggests that information is lost. The program does give some overlapping pathways in colon, but in cecum other processes are selected. In both cases no overlap with processes only selected with RF was found.

Table 5. Biological processes in the colon dataset selected by Metacore, ErmineJ and GSEA.

A: Metacore and GSEA

T-test*	p-value	RF*	p-value	GSEA genesets†	NES	p-value
mitochondrial electron transport, NADH to ubiquinone	3E-10	**mitochondrial electron transport, NADH to ubiquinone**	6E-08	***Mitochondria***	2.09	0.00
oxidative phosphorylation	4E-09	**protein targeting to mitochondrion**	1E-07	***Electron transport chain***	1.97	0.00
organelle ATP synthesis coupled electron transport	5E-09	**mitochondrial transport**	2E-07	***Oxidative phosphorylation***	1.96	0.00
ATP synthesis coupled electron transport	5E-09	**oxidative phosphorylation**	3E-07	*Propanoate metabolism*	1.92	0.00
protein targeting to mitochondrion	1E-08	**electron transport**	3E-07	*Proteasome degradation*	1.86	0.01
mitochondrial transport	2E-08	**organelle ATP synthesis coupled electron transport**	5E-07	*Proteasome*	1.83	0.01
electron transport	2E-06	**ATP synthesis coupled electron transport**	5E-07	*Free Radical Induced Apoptosis*	1.80	0.01
regulation of carbohydrate metabolic process	2E-05	**regulation of carbohydrate metabolic process**	2E-05	*Butanoate metabolism*	1.79	0.00
muscle filament sliding	1E-04	**coenzyme metabolic process**	2E-05	***Tricarboxilic acid cycle***	1.78	0.00
coenzyme metabolic process	1E-04	**energy derivation by oxidation of organic compounds**	1E-04	*Programmed cell death*	1.77	0.00
regulation of insulin secretion	2E-04	**regulation of insulin secretion**	1E-04	*Valine leucine and isoleucine degradation*	1.75	0.02
main pathways of carbohydrate metabolic process	2E-04	cofactor metabolic process	2E-04			
biopolymer catabolic process	2E-04	response to inorganic substance	2E-04			
response to copper ion	2E-04	**response to copper ion**	2E-04			
nucleosome assembly	3E-04	**nucleosome assembly**	3E-04			
feeding behavior	4E-04	**insulin secretion**	3E-04			
insulin secretion	4E-04	regulation of secretion	4E-04			
chromatin assembly	4E-04	aerobic respiration	4E-04			
energy derivation by oxidation of organic compounds	4E-04	response to metal ion	6E-04			
monocarboxylic acid metabolic process	6E-04	**main pathways of carbohydrate metabolic process**	8E-04			
response to toxin	9E-04	**chromatin assembly or disassembly**	8E-04			
		response to toxin	8E-04			
		positive regulation of pseudopodium formation	8E-04			
		regulation of hormone secretion	9E-04			
		peptide hormone secretion	9E-04			
		peptide secretion	9E-04			
		protein targeting	9E-04			

*Gene subsets of t-test and RF were used as input for Metacore. Overlapping processes between the two genesets (t-test and RF) are presented in bold.

† For GSEA the whole dataset was used, only the genesets compiled from publicly available databases are included.

B: ErmineJ and GSEA

T-test*	p-value	RF*	p-value	GSEA genesets †	NES	p-value
mitochondrial electron transport, NADH to ubiquinone	5E-18	**mitochondrial electron transport, NADH to ubiquinone**	2E-14	***Mitochondria***	2.09	0.00
ATP synthesis coupled electron transport (sensu Eukaryota)	3E-15	**ATP synthesis coupled electron transport (sensu Eukaryota)**	6E-12	***Electron transport chain***	1.97	0.00
protein biosynthesis	5E-14	**aerobic respiration**	2E-11	***Oxidative phosphorylation***	1.96	0.00
macromolecule biosynthesis	1E-13	**tricarboxylic acid cycle**	7E-09	*Propanoate metabolism*	1.92	0.00
aerobic respiration	2E-11	**protein targeting to mitochondrion**	1E-08	*Proteasome degradation*	1.86	0.01
tricarboxylic acid cycle	2E-10	**main pathways of carbohydrate metabolism**	5E-08	*Proteasome*	1.83	0.01
main pathways of carbohydrate metabolism	4E-09	**acetyl-CoA catabolism**	2E-07	*Free Radical Induced Apoptosis*	1.80	0.01
acetyl-CoA catabolism	7E-09	**protein biosynthesis**	4E-07	*Butanoate metabolism*	1.79	0.00
protein targeting to mitochondrion	1E-08	**generation of precursor metabolites and energy**	3E-06	***Tricarboxilic acid cycle***	1.78	0.00
generation of precursor metabolites and energy	1E-07	**proton transport**	6E-06	*Programmed cell death*	1.77	0.00
proton transport	5E-07	**macromolecule biosynthesis**	6E-06	*Valine leucine and isoleucine degradation*	1.75	0.02
pyruvate metabolism	2E-06	**oxidative phosphorylation**	2E-05			
hexose biosynthesis	5E-06	**pyruvate metabolism**	2E-05			
oxidative phosphorylation	2E-05	**DNA fragmentation during apoptosis**	4E-05			
hydrogen transport	2E-05	**ATP biosynthesis**	6E-05			
establishment and/or maintenance of chromatin architecture	2E-05	cellular respiration	7E-05			
DNA fragmentation during apoptosis	4E-05	**DNA catabolism**	7E-05			
fatty acid beta-oxidation	5E-05	disassembly of cell structures during apoptosis	7E-05			
ATP biosynthesis	6E-05	**ATP synthesis coupled proton transport**	7E-05			
DNA catabolism	7E-05	**coenzyme metabolism**	8E-05			
disassembly of cell structures during apoptosis	7E-05	hexose biosynthesis	9E-05			
ATP synthesis coupled proton transport	7E-05	**protein secretion**	9E-05			
protein secretion	9E-05	tricarboxylic acid cycle intermediate metabolism	1E-04			
cellular metabolism	1E-04	apoptotic nuclear changes	1E-04			
apoptotic nuclear changes	1E-04	sensory perception of sound	2E-04			
DNA packaging	1E-04	hydrogen transport	2E-04			
sensory perception of sound	2E-04	acyl-CoA metabolism	2E-04			
acyl-CoA metabolism	2E-04	purine ribonucleoside triphosphate biosynthesis	2E-04			

purine ribonucleoside triphosphate biosynthesis	2E-04	**electron transport**	3E-04
electron transport	3E-04	carbohydrate metabolism	9E-04
protein targeting	3E-04		
fatty acid oxidation	4E-04		
protein amino acid deacetylation	5E-04		
fatty acid metabolism	5E-04		
establishment of protein localization	7E-04		
coenzyme metabolism	8E-04		
oxygen and reactive oxygen species metabolism	9E-04		

* Gene subsets of t-test and RF were used as input for ErmineJ. Overlapping processes between the two genesets (t-test and RF) are presented in bold.
† For GSEA the whole dataset was used, only the genesets compiled from publicly available databases are included.

Table 6: Biological processes in the cecum dataset selected by Metacore, ErmineJ and GSEA.

A: Metacore and GSEA

T-test *	p-value	RF*	p-value	GSEA genesets†	NES	p-value
feeding behavior	2E-08	**feeding behavior**	6E-10	*Cell cycle regulator*	1.75	0.01
regulation of insulin secretion	4E-08	leading edge cell differentiation	1E-08	*Cholesteron biosynthesis*	1.67	0.02
regulation of carbohydrate metabolic process	3E-07	**regulation of insulin secretion**	3E-08	*cell proliferation*	1.66	0.00
insulin secretion	9E-07	**eating behavior**	9E-08	*Interleukin 10 pathway*	1.61	0.03
regulation of hormone secretion	3E-06	**regulation of carbohydrate metabolic process**	3E-07	*Caspase cascade*	1.56	0.04
peptide hormone secretion	3E-06	**insulin secretion**	7E-07	*Proliferation*	1.56	0.00
peptide secretion	3E-06	**regulation of hormone secretion**	2E-06			
eating behavior	6E-06	**peptide hormone secretion**	2E-06			
peptide transport	9E-06	**peptide secretion**	2E-06			
regulation of lipid metabolic process	5E-05	**peptide transport**	7E-06			
hormone secretion	5E-05	epithelial cell differentiation	2E-05			
regulation of angiogenesis	2E-04	**regulation of secretion**	3E-05			
regulation of secretion	4E-04	hormone secretion	4E-05			
generation of a signal involved in cell-cell signaling	5E-04	**regulation of lipid metabolic process**	4E-05			
monocarboxylic acid transport	5E-04	morphogenesis of an epithelium	1E-04			
		cellular defense response	3E-04			
		generation of a signal involved in cell-cell signaling	4E-04			

*Gene subsets of t-test and RF were used as input for Metacore. Overlapping processes between the two genesets (t-test and RF) are presented in bold.
† For GSEA the whole dataset was used, only the genesets compiled from publicly available databases are included.

B: Ermine J and GSEA

| T-test*	p-value	RF*	p-value	GSEA genesets†	NES	p-value
digestion	2E-06	**cellular defense response**	1E-06	*Cell cycle regulator*	1.75	0.01
regulation of angiogenesis	6E-06	epithelial cell differentiation	1E-05	*Cholesteron biosynthesis*	1.67	0.02
cellular defense response	4E-05	**oxygen and reactive oxygen species metabolism**	3E-05	*cell proliferation*	1.66	0.00
muscle development	6E-05	neuron migration	3E-05	*Interleukin 10 pathway*	1.61	0.03
wound healing	3E-04	T cell activation	4E-05	*Caspase cascade*	1.56	0.04
actin cytoskeleton organization and biogenesis	4E-04	response to wounding	5E-05	*Proliferation*	1.56	0.00
oxygen and reactive oxygen species metabolism	5E-04	**digestion**	5E-05			
tissue development	6E-04	defense response	1E-04			
regulation of cell differentiation	9E-04	cell migration	2E-04			
		wound healing	3E-04			
		cell motility	7E-04			
		regulation of cell differentiation	8E-04			
		neurogenesis	9E-04			

*Gene subsets of t-test and RF were used as input for ErmineJ. Overlapping processes between the two genesets (t-test and RF) are presented in bold.

† For GSEA the whole dataset was used, only the genesets compiled from publicly available databases are included.

Discussion

We described a framework for physiological interpretation of gene expression data. This framework (see BOX 1) consists of the following steps: genes are first ranked, the relevant genes are selected, the selected genes are grouped according to their expression profile and then biologically interpreted. The considerations underlying the different steps are illustrated using two real gene expression data sets. We show several features of Random Forest (RF) that should be part of any data-analysis framework. These are 1) all genes in the dataset are included in the analysis, 2) interaction between genes is taken into account and, 3) a well-defined gene set can be selected using an objective threshold.

BOX 1. A framework for identification of physiological responses in microarray based gene expression studies. The framework is composed of the following steps: Gene ranking, gene selection, gene grouping and biological interpretation. Essential features of the data-analysis framework are that 1) all genes (annotated and non-annotated) in the dataset are included in the analysis, 2) interaction between genes is taken into account and, 3) an objective threshold is used for selection of a well-defined gene set. Random forest has these features. Gene grouping can provide information on new targets and add information above pathway analysis. Despite loss of information due to incomplete annotation of the complete dataset, Gene set enrichment analysis can provide additional information on related genes with small differences. (For full color figure, see page 171).
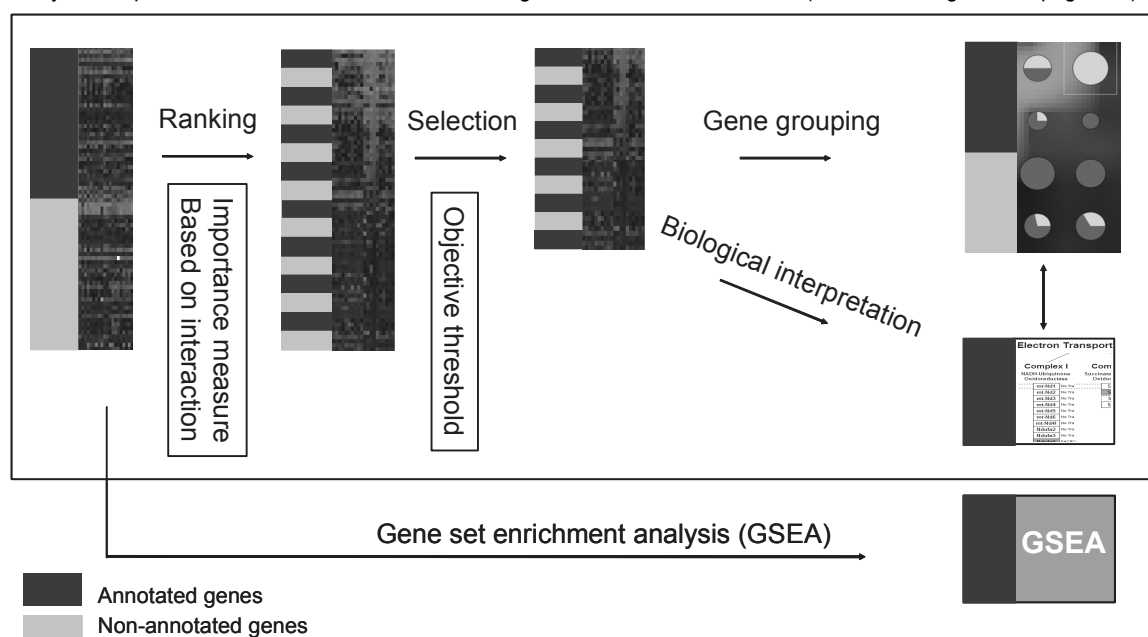
For human, mouse and rat whole genome arrays, the number of annotated genes is less than half of the genes present on the array. Consequently, analysis only based on functional annotation and co-occurrence in gene sets filters out half of the information present in the microarray dataset. Well studied biological processes are better represented in pathway databases[22]. Therefore, conclusions obtained from data analysis based only on pathway programs are biased towards the well annotated biological processes. By including all genes from a whole genome dataset, it is possible to find genes or processes less defined in databases, but which could be attractive new targets for drug development or nutritional intervention. For both colon and cecum, genes exclusively selected within one SOM profile belonged to the same biological process: cellular

component organization and biogenesis (colon) and immune response (cecum) respectively. As only a few genes within these profiles were GO-annotated, these processes were not selected by the different pathways programs. By literature and database search we could clearly identify these genes as part of this process.

A major strength of whole genome microarray studies is that the expression levels of all genes are displayed, allowing for identification of gene-gene interactions. RF was chosen to rank genes, as its measure of importance takes possible interactions between genes into account. Compared with the results obtained by t-test, RF selected genes with main effects but additionally was able to capture weak effects. In studies with small gene expression changes which are independently not significant, but occurring in one group may be of large relevance, this is an advantage. For example, it enables identification of possible side effects in drug studies, or expected subtle differences in nutritional studies. In our study, application of RF in combination with SOM indeed showed enriched profiles containing mainly genes selected exclusively by RF and not by t-test. Genes within these profiles are therefore contributing by gene-gene interactions.

By applying a permutation test we defined a threshold for RF to select genes in an objective way. Comparison of different runs showed that the most important genes were consistently selected. However, selection of genes ranked closely above the threshold varied between different runs. We chose to include genes that were additionally selected over different runs in the total selected gene set. By including genes selected additionally by different runs there is a chance that more false positives were included in the selection. If we would have chosen to select the set of genes that overlapped in all runs, we might discard truly relevant genes (false negatives). We reasoned that the increased information available for pathway analysis outweighed the potential disadvantage of including some noise, especially since in dietary studies gene expression changes of interest are usually small. Furthermore, the results show that the number of additionally selected genes decreased rapidly for each additional run. As there was large overlap, it is less likely that many of the additionally selected genes were noise. Thus, within this framework, RF is a useful tool to select a well-defined set of genes for further interpretation.

SOM was applied to find groups of genes with similar gene expression profiles. Other approaches to find gene groups, such as hierarchical clustering, can be used with the same objective[34]. SOM has the advantage that clustering is organized in profile-groups and therefore provides a more ordered output than that of other cluster programs. While individual genes may have small gene expression differences, groups of similarly behaving genes can be biologically significant. When SOM analysis is applied to whole genome datasets, unrelated data will also produce clusters, without any physiological relevance[34]. This can be overcome by selecting a subset of genes and to examine whether biological valid clusters are obtained. For both colon and cecum, genes selected by RF and analyzed by SOM provided profiles consisting of genes with similar biological function. In the colon dataset, a SOM profile consisted of genes belonging to the same GO-process, and genes with poorly identified functions. This could be a starting point to identify possible biological function of the non-identified genes. Using SOM within this framework can provide information on genes with unknown function and help to identify

biological processes not captured by pathway analyses. Therefore SOM is a useful tool for identification of biological processes in addition to pathway analysis.

The pathway analysis based on the subset of genes obtained by RF and t-test shows overlap for the selected processes, however different processes were additionally obtained by RF.
Remarkably, GSEA only returned a few gene sets connected to public databases that were significantly enriched in colon or in cecum. The small number of processes identified by GSEA analysis suggests that information is lost. On the other hand, GSEA did provide biological processes not found in the other pathway programs. Although only a few processes were found by GSEA, these are worth exploring as these may consist of related genes with small differences. Thus, in the context of the framework discussed in this paper, GSEA may additionally be applied.

The advantage of this framework is that different methods can be applied at different steps, for example a Bayesian method, depending on the aim and preferences of the researcher. This study used two real datasets with subtle gene expression changes and showed that RF can be used to extract a biologically meaningful group of genes, such as the set of immune response genes in the cecum dataset that would be discarded with univariate tests such as the t-test. Previous studies on simulation datasets by Lunetta *et al*[26] showed that RF outperformed other univariate methods. RF is not the only method available, however it has advantages to use RF within this framework to rank and select genes. As mentioned, it returns an importance factor for each gene (I_m) in which gene-gene interactions are taken into account. Based on this I_m, we showed an approach which can be used to define an objective threshold for selection of genes.
Besides two classes RF can also be applied to multi-class problems. Furthermore, free software is available for RF whereby only a few parameters need to be defined. Also, users can easily obtain a gene list for further interpretation without the need to understand the finer details of the method thoroughly. Therefore, within this framework RF is a suitable and practical tool to rank and select genes. Combined with gene grouping by SOM and pathway programs, this framework is helpful to obtain insight in the biological processes. These physiological effects are the main focus for further confirmatory and mechanistic studies.

In conclusion, in this study we have examined the application of a framework in which all genes in a microarray dataset are analyzed. Within this framework, application of RF has the advantage that it takes gene-gene interactions in the ranking of genes into account. Also, selection of genes by an objective threshold provides a well-defined set of genes for further interpretation. Groups of genes within this set are identified by SOM-analysis. It provides in combination with pathway analyses valuable information on biological processes involved in the treatment.

Acknowledgements

References

1. Afman, L. & Muller, M. Nutrigenomics: from molecular nutrition to prevention of disease. *J Am Diet Assoc* 106, 569-576 (2006).
2. Allison, D.B., Cui, X., Page, G.P. & Sabripour, M. Microarray data analysis: from disarray to consolidation and consensus. *Nat Rev Genet* 7, 55-65 (2006).
3. Barry, W.T., Nobel, A.B. & Wright, F.A. Significance analysis of functional categories in gene expression studies: a structured permutation approach. *Bioinformatics* 21, 1943-1949 (2005).
4. Brazhnik, P., de la Fuente, A. & Mendes, P. Gene networks: how to put the function in genomics. *Trends Biotechnol* 20, 467-472 (2002).
5. Breiman, L. Fortran code for Random Forests. www.stat.berkeley.edu/user/breiman/randomforests/.
6. Breiman, L. Random Forest. *Machine Learning* 45, 5-32 (2001).
7. Breitling, R., Armengaud, P., Amtmann, A. & Herzyk, P. Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Lett* 573, 83-92 (2004).
8. Chen, J.J., Wang, S.J., Tsai, C.A. & Lin, C.J. Selection of differentially expressed genes in microarray data analysis. *Pharmacogenomics J* (2006).
9. Cox, D.R. & Hinkley, D.V. *Theoretical statistics*, (London, 1974).
10. Cran-website. http://cran.r=project.org/.
11. Curk, T*., et al.* Microarray data mining with visual programming. *Bioinformatics* 21, 396-398 (2005).
12. de Boer, V.C*., et al.* Chronic quercetin exposure affects fatty acid catabolism in rat lung. *Cell Mol Life Sci* 63, 2847-2858 (2006).
13. Diaz-Uriarte, R. & Alvarez de Andres, S. Gene selection and classification of microarray data using random forest. *BMC Bioinformatics* 7, 3 (2006).
14. Draghici, S., Sellamuthu, S. & Khatri, P. Babel's tower revisited: a universal resource for cross-referencing across annotation databases. *Bioinformatics* 22, 2934-2939 (2006).
15. Ekins, S., Nikolsky, Y., Bugrim, A., Kirillov, E. & Nikolskaya, T. Pathway mapping tools for analysis of high content data. *Methods Mol Biol* 356, 319-350 (2007).
16. Gene Ontology Consortium. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res* 32, D258-261 (2004).
17. GeneSrF: Gene selection with Random forest. http://genesrf.bioinfo.cnio.es, CNIO, Bioinfromatics unit.
18. Goeman, J.J. & Buhlmann, P. Analyzing gene expression data in terms of gene sets: methodological issues. *Bioinformatics* 23, 980-987 (2007).
19. Hochberg, Y. & Benjamini, Y. More powerful procedures for multiple significance testing. *Stat Med* 9, 811-818 (1990).
20. Huang, X*., et al.* A comparative study of discriminating human heart failure etiology using gene expression profiles. *BMC Bioinformatics* 6, 205 (2005).
21. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28, 27-30 (2000).
22. Khatri, P. & Draghici, S. Ontological analysis of gene expression data: current tools, limitations, and open problems. *Bioinformatics* 21, 3587-3595 (2005).
23. Lee, H.K., Braynen, W., Keshav, K. & Pavlidis, P. ErmineJ: tool for functional analysis of gene expression data sets. *BMC Bioinformatics* 6, 269 (2005).
24. Lee, J.W., Lee, J.B., Park, M. & Song, S.H. An extensive comparison of recent classifcation tools applied to microarray data. *Computational statistics & Data analysis* 48, 869-885 (2005).
25. Liaw, A. & Wiener, M. Classification and regression by randomforest. *R News* 2, 18-22 (2002).

26. Lunetta, K.L., Hayward, L.B., Segal, J. & Van Eerdewegh, P. Screening large-scale association study data: exploiting interactions using random forests. *BMC Genet* 5, 32 (2004).

27. Lyons-Weiler, J*., et al.* Assessing the statistical significance of the achieved classification error of classifiers constructed using serum peptide profiles, and a prescription for random sampling repeated studies for massive high-throughput genomic and proteomics studies. *Cancer Informatics* 1, 53-77 (2005).

28. Majumder, P.K*., et al.* mTOR inhibition reverses Akt-dependent prostate intraepithelial neoplasia through regulation of apoptotic and HIF-1-dependent pathways. *Nat Med* 10, 594-601 (2004).

29. Mootha, V.K*., et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 34, 267-273 (2003).

30. Norris, A.W. & Kahn, C.R. Analysis of gene expression in pathophysiological states: Balancing false discovery and false negative rates. *PNAS* 103, 649-653 (2006).

31. Parast, M.M. & Otey, C.A. Characterization of palladin, a novel protein localized to stress fibers and cell adhesions. *J Cell Biol* 150, 643-656 (2000).

32. Patsouris, D., Reddy, J.K., Muller, M. & Kersten, S. Peroxisome proliferator-activated receptor alpha mediates the effects of high-fat diet on hepatic gene expression. *Endocrinology* 147, 1508-1516 (2006).

33. Pellis, L., Franssen-van Hal, N.L., Burema, J. & Keijer, J. The intraclass correlation coefficient applied for evaluation of data correction, labeling methods, and rectal biopsy sampling in DNA microarray experiments. *Physiol Genomics* 16, 99-106 (2003).

34. Quackenbush, J. Computational analysis of microarray data. *Nat Rev Genet* 2, 418-427 (2001).

35. R development core team. R: A language and environment for statistical computing. in *http://www.R-project.org* (2004).

36. Rubin, E. Circumventing the cut-off for enrichment analysis. *Brief Bioinform* 7, 202-203 (2006).

37. Segal, E., Friedman, N., Kaminski, N., Regev, A. & Koller, D. From signatures to models: understanding cancer using microarrays. *Nat Genet* 37 Suppl, S38-45 (2005).

38. Shi, T., Seligson, D., Belldegrun, A.S., Palotie, A. & Horvath, S. Tumor classification by tissue microarray profiling: random forest clustering applied to renal cell carcinoma. *Mod Pathol* 18, 547-557 (2005).

39. Slonim, D.K. From patterns to pathways: gene expression data analysis comes of age. *Nat Genet* 32 Suppl, 502-508 (2002).

40. Smyth, G.K., Yang, Y.H. & Speed, T. Statistical issues in cDNA microarray data analysis. *Methods Mol Biol* 224, 111-136 (2003).

41. Strobl, C., Boulesteix, A.L., Zeileis, A. & Hothorn, T. Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics* 8, 25 (2007).

42. Subramanian, A*., et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102, 15545-15550 (2005).

43. Tamayo, P*., et al.* Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proc Natl Acad Sci U S A* 96, 2907-2912 (1999).

44. Tusher, V.G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 98, 5116-5121 (2001).

45. Valafar, F. Pattern recognition techniques in microarray data analysis: a survey. *Ann N Y Acad Sci* 980, 41-64 (2002).

46. Verducci, J.S*., et al.* Microarray analysis of gene expression: considerations in data mining and statistical treatment. *Physiol Genomics* 25, 355-363 (2006).

47. Werner, T. Regulatory networks: linking microarray data to systems biology. *Mech Ageing Dev* 128, 168-172 (2007).

48. Yoon, S., Yang, Y., Choi, J. & Seong, J. Large scale data mining approach for gene-specific standardization of microarray gene expression data. *Bioinformatics* 22, 2898-2904 (2006).

Chapter 6

Impaired intestinal barrier function by dietary fructo-oligosaccharides (FOS) in rats is associated with increased mitochondrial gene expression

Wendy Rodenburg[1,2,3]
Jaap Keijer[1,2]
Evelien Kramer[1,2]
Carolien Vink[1, 3]
Roelof van der Meer[3,4]
Ingeborg Bovee-Oudenhoven[1,3]

[1] TI Food and Nutrition, Wageningen, The Netherlands

[2] RIKILT - Institute of Food Safety, Wageningen, The Netherlands

[3] NIZO Food Research, Ede, The Netherlands

[4] Nutrigenomics Consortium, TI Food and Nutrition, Wageningen, The Netherlands

Abstract

Background

Dietary non-digestible carbohydrates stimulate the gut microflora and are therefore presumed to improve host resistance to intestinal infections. However, several strictly controlled rat infection studies showed that non-digestible fructo-oligosaccharides (FOS) increase, rather than decrease, translocation of Salmonella towards extra-intestinal sites. In addition, it was shown that FOS increases intestinal permeability already before infection. The mechanism responsible for this adverse effect of FOS is unclear. Possible explanations are altered mucosal integrity due to changes in tight junctions or changes in expression of defense molecules such as antimicrobials and mucins. To examine the mechanisms underlying weakening of the intestinal barrier by FOS, a controlled dietary intervention study was performed. Two groups of 12 rats were adapted to a diet with or without FOS. mRNA was collected from colonic mucosa and changes in gene expression were assessed for each individual rat using Agilent rat whole genome microarrays.

Results

Among the 997 FOS induced genes we observed less mucosal integrity related genes than expected with the clear permeability changes. FOS did not induce changes in tight junction genes and only 8 genes related to mucosal defense were induced by FOS. These small effects are unlikely the cause for the clear increase in intestinal permeability that is observed. FOS significantly increased expression of 177 mitochondria-related genes. More specifically, induced expression of genes involved in all 5 OXPHOS complexes and the TCA cycle was observed. These results indicate that dietary FOS influences intestinal mucosal energy metabolism. Furthermore, increased expression of 113 genes related to protein turnover, including proteasome genes, ribosomal genes and protein maturation related genes was seen. FOS induced increase in the peptide hormone proglucagon gene, in agreement with previous studies, as well as three other peptide hormone genes; peptide YY, pancreatic polypeptide and cholecystokinin.

Conclusion

We conclude that altered energy metabolism underly barrier function disruption due to FOS feeding in rats.

Introduction

Non-digestible carbohydrates like fructo-oligosaccharides (FOS) stimulate the gut microflora and are therefore presumed to improve host resistance to intestinal infections. For this reason non-digestible carbohydrates are added to a growing list of products, including baby-formula, bread, dairy products. Many studies, including our own, showed that non-digestible carbohydrates indeed affect intestinal microflora composition[9,28,76]. However, there is little evidence that these non-digestible carbohydrates strengthen intestinal resistance to infection and gut barrier function. For this reason, several strictly controlled rat infection studies were previously performed at our lab. These studies consistently showed that the non-digestible carbohydrates inulin, lactulose and FOS increase translocation of Salmonella to extra-intestinal organs[74,76,77]. A dose-dependent increase in Salmonella translocation was observed in FOS-fed rats[76]. Stimulation of Salmonella translocation by dietary FOS was reflected in transcriptional changes in colon. Genes involved in antimicrobial defense, immune response and inflammation were induced by Salmonella infection of rats on a control diet and further upregulated in Salmonella infected rats on a FOS diet [61]. Moreover, intestinal barrier parameters were already affected by FOS before infection. In particular intestinal permeability was increased by FOS before Salmonella challenge. Also, luminal cytotoxicity and faecal mucin excretion were increased in FOS-fed rats and may indicate mucosal irritation[77].

The intestinal barrier is mainly formed by the mucosal epithelial lining. Disturbed barrier function can be monitored by measurement of epithelial permeability using inert permeability markers such as different kind of sugars or CrEDTA[4,82]. It has been shown that increased transport of large molecules or antigens due to increased permeability may initiate inflammation[7]. Several mechanisms have been implicated in the mucosal barrier. Tight junctions tightly connect the epithelial cells and regulate paracellular transport of fluids, electrolytes and small compounds[44]. Modification or cellular translocation of the tight junction molecules ZO1 and several claudins have indeed been observed in inflammatory bowel disease (IBD) and chronic non-steroidal anti-inflammatory drugs (NSAIDs) use, both characterized by increased intestinal permeability[46,85].

In addition to changes in epithelial tight junctions, balance between apoptosis and proliferation, or regeneration, is also a major determinant of an intact mucosal epithelial lining[54]. Increased apoptosis can induce epithelial leakage as shown in colonic epithelial cell lines and in intestinal biopsy specimens of IBD patients[68].

Apart from tight junctions and apoptosis, secretory products of intestinal epithelial cells are known to play a role in mucosal barrier. The different intestinal epithelial cells; enterocytes, goblet cells, Paneth cells and enteroendocrine cells, are equipped with defense mechanisms. For instance Paneth cells in the crypt base produce antimicrobials, such as defensins, lysozyme and Pla2g2a[83], to regulate and restrict the bacterial load in the gut lumen. Goblet cells produce mucins to form a mucus layer, this layer functions as a physical barrier that protects the epithelial cells from harmful compounds[73]. Furthermore, the immune system plays an important role in mucosal integrity[49]. Whether other processes are involved and the relative importance of these

mechanisms for intestinal barrier integrity is not known. Also it is not known if any of the above mentioned mechanisms or others are responsible for the effect of FOS on gut permeability. Therefore, we analyzed colonic gene expression changes in individual rats fed a 6% FOS diet for 2 weeks. Analyzing FOS induced gene expression using whole genome microarrays allowed us to not only focus on the above mentioned mechanisms, but to obtain an unbiased view on processes affected by dietary FOS. This facilitates identification of genes and processes currently unknown to be related to barrier function. The colon had our main interest as FOS and other non-digestible carbohydrates are exclusively fermented in the distal gut in humans[2] and rats[76]. We aimed to identify the in vivo biological mechanisms involved in FOS-induced weakening of the barrier in rats.

Materials and Methods

Animals and diet

The animal welfare committee of Wageningen University (Wageningen, the Netherlands) approved the experimental protocol. Specific pathogen-free male outbred Wister rats (8 weeks old, mean body weight of 253 g; n=36 in total), were housed individually in metabolic cages. All animals were kept in a temperature (22-24 C) and humidity (50-60%) controlled room with a 12 h light/dark cycle (lights on from 6 AM to 6 PM). Rats (two dietary groups, n=18 each) were fed restricted quantities (14 g/day) of a purified diet during the entire experimental period. Restricted food intake was necessary to prevent differences in food consumption and hence differences in vitamin and mineral intake as observed earlier in FOS interventions[76]. The diet contained (per kg) 200 g acid casein, 502 g glucose, 160 g palm oil, 40 g corn oil, 20 g cellulose, 35 g mineral mix (without calcium) and 10 g vitamin mix according to AIN93 recommendations[59]. Diets contained 20 g/kg cellulose at least and were supplemented with either 60 g/kg FOS (purity 93%; Raftilose P95, Orafti, Tienen, Belgium) or additional 60 g/kg cellulose as described earlier[76]. Diets were low in calcium (20 mmol $CaHPO_4.2H_2O$/kg) and high in fat content (200 g fat/kg) to mimic the composition of a Western human diet. Demineralized drinking water was supplied *ad libitum.*

To follow intestinal permeability, 6 of the 18 rats of each dietary group received their diet supplemented with the intestinal permeability marker chromium ethylenediamine-tetraacetic acid (CrEDTA). The CrEDTA solution added to the diet was prepared as described elsewhere[77]. After feeding the diets for 16 days, rats were killed by carbon dioxide inhalation. Rats fed diets containing the permeability marker CrEDTA (n=6 per diet group) were not included in the gene expression study, to exclude possible interaction of CrEDTA on colonic gene expression. From the remaining 12 rats per dietary group, the colon was taken out, longitudinally opened and colonic contents were removed by a quick rinse in 154 mM KCl. Colonic mucosa was scraped off using a spatula. Scrapings were immediately frozen in liquid nitrogen and stored at -80 C. The scrapings were homogenized in liquid N_2 using a mortar and pestle cooled with liquid N_2 (Fisher Emergo, Landsmeer, The Netherlands). One third of the pulverized samples was used for protein determination and the remaining part for RNA isolation.

Analysis of urine samples

Total 24 h urine samples were collected on days 14 and 15 from rats fed the control and FOS diet (n=6 each) supplemented with CrEDTA. Urines were preserved by adding oxytetracycline (1 mg) to the urine collection vessels of the metabolic cages, and analyzed for the intestinal permeability marker CrEDTA as described elsewhere[77]. CrEDTA data were analyzed using the Student's t-test (two-sided) using Prism 4 (GraphPad software Inc., San Diego, CA).

RNA isolation

Total RNA was isolated from colon scraping homogenates using TRIzol reagent (Invitrogen, San Diego, CA) according to the manufacturer's instructions. Total RNA was purified using Rneasy columns (Qiagen, Chatsworth, CA). Absence of RNA degradation was checked on a 1% TBE/agarose gel after 1 hour incubation at 37°C. RNA purity and concentration were measured with the Nanodrop (Isogen Life Science, Maarssen, The Netherlands). OD A_{260}/A_{280} ratios were all between 2.08 and 2.10 indicating RNA of high purity.

Analysis of mRNA expression by Oligo Arrays

For microarray hybridization, RNA of each individual animal was labeled with Cy5. A standard reference sample, consisting of a pool of all colonic RNA was labeled with Cy3. For each oligo array, 35 μg of total RNA was used to make Cy5 or Cy3 labeled cDNA. Total RNA was mixed with 4 μg T21 primer, heated at 65°C for 3 min (RNA denaturation) followed by 25°C for 10 min (primer annealing). cDNA was synthesized by adding 5x first strand buffer (Invitrogen), 10 mM DTT, 0.5mM dATP, 0.5 mM dGTP, 0.5 mM dTTP, 0.04 mM dCTP, 0.04 mM Cy5-dCTP or Cy3-dCTP, 1.2U RnaseOUT and 6U SuperScript II Reverse Transcriptase to a total volume of 62.5 μL. The reaction was incubated at 42°C for 2 h. Purification, precipitation and denaturation of the labeled cDNA were performed as previously described[80].

Each labeled cDNA sample was individually hybridized on the 44K rat whole genome Agilent array (G4131A, Agilent Technologies, Inc. Santa Clara, CA) consisting of 44290 60-mer rat oligonucleotide probes, including ~3000 control spots. The Cy5 labeled cDNAs of the individual rats were mixed 1:1 with the Cy3 labeled reference cDNA, mixed with 2x hybridization buffer (Agilent Technologies) and 10x control targets (Agilent Technologies) and hybridized for 17 hours at 60°C in Agilent hybridization chambers in an Agilent hybridization oven rotating at 4 rpm. After hybridization the arrays were washed with an SSPE wash procedure (Agilent Technologies) and scanned with a Scanarray Express HT scanner (Perkin Elmer).

Data analyses and functional interpretation of microarray data

Spot intensities were quantified using ArrayVision 8.0 (GE Healthcare life sciences). Median density values and background values of each spot were extracted for both the experimental samples (Cy5) and the reference samples (Cy3). Subsequently, quality control was performed for each microarray using both LimmaGUI package in R from Bioconductor and Microsoft Excel. One array in the dietary FOS group did not pass the quality control based on MA plot and signal intensity distribution [1]. Therefore, the dataset contained 23 arrays in total. Data was exported into GeneMaths XT (Applied Maths, Sint-Martens-Latem, Belgium) for background correction and normalization. We discarded spots with an average intensity, over all arrays, of Cy5 lower than

1.5-fold above average background. Then, the Cy5 intensities were normalized against the Cy3 reference as described previously[53]. The data have been deposited in NCBIs Gene Expression Omnibus[52] and are accessible through GEO Series accession number GSE5943. The complete dataset is available as supplemental table at www.foodbioactives.nl. Fold changes calculations were performed in Microsoft Excel, fold change equals ratio FOS/control in the case of increase or equals -1/ratio in the case of decrease. For statistical identification of differentially regulated probes between the control and FOS group we used two complementary tests, the often used t-test and Random Forest (RF). T-test was performed in GeneMaths XT, the generated p-values were used to obtain insight into significantly affected genes. To correct for multiple testing we used FDR-adjusted p-values (GeneMaths XT), so called q-values[6]. For t-test we choose a stringent threshold of $p<0.001$. The corresponding q-value was 0.035, meaning that 3.5% of the genes selected by this p-value could be false positive. The t-test tests each gene independently and therefore will miss genes that have no main effect but are related to the treatment in gene-gene interaction[56]. We therefore used RF, available as R-package[17,57], as a complementary method as that method includes genes that in gene-gene interaction are related to treatment besides including genes with a main treatment effect. RF was recently successfully used in several microarray studies[20,40]. The method provides an importance index for each gene. This value is dependent on the main treatment effect of a gene. In addition, gene-gene interaction related to the treatment increases importance index value of genes[40]. For RF we defined a threshold where the importance index of each gene in the real dataset exceeded the importance index of genes obtained from analysis of 100 randomly permuted datasets, using randomly assigned class labels FOS or control. This indicates that these genes are truly related to the treatment[41] (detailed method described in Chapter 5). We included the genes that were selected by the t-test threshold and the genes selected by the RF threshold. These genes were considered significantly changed by FOS.

To interpret functional changes in the dataset, we applied two pathway analysis programs, Metacore and GSEA, with different complementary pathway-classification properties. Pathway analysis of the selected genes was performed using MetaCore (GeneGo Inc, St. Joseph, MI). We used classification based on GO-term and classification based on GeneGo annotation. The GeneGo annotation database is a curated database of gene networks based on several databases (KEGG, GO) and scientific literature[24]. We also performed a pathway analysis with GSEA (Broad Institute), a method that does not require preselection of genes by a statistical threshold but uses the whole dataset. GSEA is thoroughly described by Subramanian et al[72]. This method prevents possible selection bias[14,63]. We used the c2 functional genesets based on publicly available and curated databases (GenMapp, Biocart and SigmaAldrich). Only processes with 5-500 genes were taken into account. Agilent gene annotation version 20060331 was used for both programs. We selected pathways with $p<0.001$ in metacore and q-value <0.25 in GSEA, in accordance with the recommendation of the GSEA developers[72].

Since only about 30% of the genes on the whole genome array were recognized in both pathway programs, we manually supplemented the significantly enriched biological processes with non-annotated genes from the selected gene-set using biological databases (BIOcarta, SOURCE, GenMAPP, KEGG) and scientific literature. As processes overlap, we bundled some processes and renamed them.

Analysis of mRNA expression by Real-time Quantitative RT-PCR

Real-time Quantitative RT-PCR (Q-PCR) was performed on individual samples (n=12 per group). 1 μg of RNA of all individual samples was used for cDNA synthesis using the iScript cDNA synthesis kit of Bio-Rad Laboratories (Veenendaal, The Netherlands). Real-time reactions were performed by means of the iQ SYBR Green Supermix of Bio-Rad using the MyIQ single-color real-time PCR detection system (Bio-Rad). Each reaction (25 μl) contained 12.5 μl IQ SYBR green supermix, 1 μl forward primer (10 μM), 1 μl reverse primer (10 μM), 8.5 μl RNase-free water and 2 μl diluted cDNA. The following cycles were performed: 1x 3 min at 95°C, 40 amplification cycles (40x 10 s 95°C, 45 s 60°C), 1x 1 min 95° C, 1x 1 min 62°C and a melting curve (80x 10 s 55°C with an increase of 0.5°C per 10 s). A negative control without cDNA template was run with every assay. The optimal melting point of dsDNA (Tm) and the efficiency of the reaction were optimized beforehand. Data were normalized against the reference genes Pleckstrin homology domain containing, family A member 6 (*Plekha6*), Nucleoporin 37 (*Nup37*) and β-actin. *Plekha6* and *Nup37* were chosen because our microarray data showed equal expression levels for all microarrays, and β-actin was chosen because it is a well accepted reference gene. Primers were designed using Beacon designer 7.00 (Premier Biosoft International, Palo Alto, CA). For primer sequences see supplemental table 6.1. A standard curve for all genes, including reference genes, was generated using serial dilutions of a pooled sample (cDNA from all reactions). mRNA levels were determined using delta CT method (IQ5 software version 2.0, Bio-Rad version). Analysis of all individual samples was performed in duplicate. Data were analyzed using Student's t-test (two-sided) using Prism 4. Differences were considered statistically significant when $p < 0.05$.

Protein determination

Mucosal scrapings of individual rats or pools of all rats per group (n=12) were lysed in a buffer containing 0.125 M TrisHCl pH 6.8, 2% SDS and 20% glycerol. Protein concentrations were determined using DC protein assay kit (Bio-rad Laboratories, Veenendaal, the Netherlands). All samples were boiled in sample buffer (0.125 M TrisHCl pH 6.8; 2% SDS; 20% glycerol; 2% β-mercaptoethanol; 0.04% coomassie briljant blue), and separated by 14% SDS-PAGE. The proteins were transferred onto a nitrocellulose membrane. Immunoblot analysis was performed with a 1:2000 dilution monoclonal antibodies against OXPHOS complexes, Complex IV subunit II (COXII) monoclonal antibody (MS601, Mitosciences, Eugene, OR, USA) or with a 1:200 dilution of monoclonal anti-GLP1 antibody (Abcam (Ab23468), Cambridge, UK) by incubation in 2.5% protifar/TBS-T for 1½ hr at RT. After incubation, blots were washed in TBS-T and incubated 1 hr at RT with a 1:2000 dilution horseradish peroxidase-conjugated anti-mouse IgG (7076, Cell Signaling, Danvers, MA, USA) for detection GLP1 or a 1:2500 dilution of horseradish peroxidase-conjugated anti-mouse IgG (W4021, Promega, Madison, WI, USA) for detection of Complex IV. The signal was detected using an enhanced chemiluminescence detection system (GE Healthcare, The Netherlands) according to the protocol of the supplier. After washing the membranes thoroughly with TBS-T, they were subsequently incubated with the monoclonal anti-Actin (1:100 dilution, Santa Cruz, sc-1615) and HRP anti-goat (1:10000 dilution, Promega. V8051). The intensities of GLP1, COXII and Actin signals on the autoradiography films were

quantified using geldoc (Bio-Rad). GLP1 and COXII quantities were normalized to actin to correct for loading differences.

Results

Food intake, body weight gain, and intestinal permeability

Rats on the control diet and FOS diet showed no significant difference in body weight gain. Both dietary groups consumed the provided 14 grams of diet per day as intended and thus had a similar dietary CrEDTA intake. Intestinal permeability was examined by measurement of CrEDTA excretion in urine and showed that FOS fed rats had increased urinary excretion of this inert permeability marker as compared to the control group (figure 1).
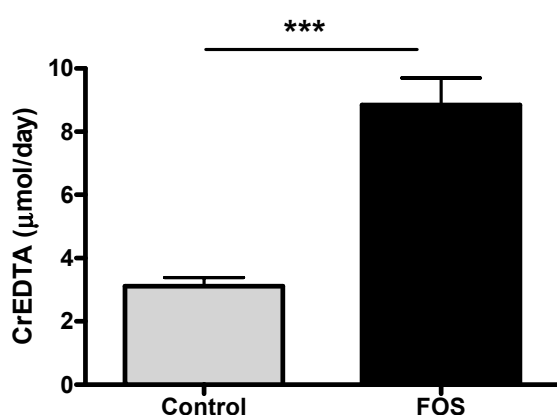
Figure 1. Average urinary CrEDTA excretion in the control and FOS group. Daily dietary CrEDTA intake was 54µmol. Urines were collected at days 14 and 15. Results are expressed as mean ± SEM (n=6 per diet group). The FOS groups significantly differed from the control group (***$p<0.001$).

Gene expression profile

28180 probes on the array had an expression value of 1.5 times above background. Of these 123 were induced and 56 were reduced more than 1.5 fold in FOS fed rats compared with rats fed the control diet. Only 19 probes were induced more than 2-fold. While gene expression changes induced by FOS supplementation were small, there were many significantly changed probes in the dataset: 803 had a p-value < 0.001, and 231 probes had a FDR derived q-value<0.01. This indicated that the gene expression response of colon mucosa to FOS was small in magnitude, but highly significant. As each probe signal is tested independently in t-test and multiple testing increases the risk for finding false positives, we also applied Random Forest[20,40]. RF ranks genes based on an importance value taking main treatment effect as well as gene-gene interaction into account[40]. Using RF, we identified 935 probes with an importance value above threshold. 629 were selected by both the t-test and RF. To prevent loss of information, we used all genes selected by t-test and RF. This resulted in a list of 1109 probes. 112 corresponding genes were listed more than once. Therefore, duplicates with the highest p-values were removed, resulting in 997 unique genes regulated by FOS

Gene selection by t-test or RF alone substantially overlapped and did result in a highly comparable outcome in pathway analysis, showing that the main effects extracted by both methods were similar. However, the individual gene selection differed slightly between both methods. We choose to include all genes selected by either method.

Effect of FOS on the expression of barrier associated genes

As FOS affects the mucosal barrier, we specifically analyzed genes that are known or assumed to play a role in barrier function. These include tight junction genes like *Zo1*, occludin and claudin, cell turnover/apoptosis genes such as caspases, *Bak, Bcl2*, and mucosal defense genes such as defensins, lipocalin, toll like receptors and IgA (supplemental table 6.2). Tight junction related genes were not affected by FOS. Several apoptosis (for example *Bax, DNase1, Pdcd 6* and *8*) and mucosal defense genes (for example phospholipase A2 and trefoil factor 1 & 3) were increased by FOS (supplemental table 6.2). However, no FOS effect was found on other mucosal defense genes like IgA, Mucin 2 & 3, defensins, lipocalin, calprotectin, and most toll-like receptors. In addition, some markers of apoptosis were slightly affected (*Bak,* Caspase 7), while most (including *Apaf,* caspase 9, caspase 3, caspase 2*, Bcl2* and *Bad*) were not affected by FOS.

Genes most prominently affected by FOS

We examined the top 10 of genes most affected by the FOS diet (table 1). We choose the genes most prioritized by RF. These genes were characterized by extremely low p-values and relatively high fold changes. The genes were related to nutrient homeostasis (proglucagon), energy metabolism (NADH dehydrogenases (*Ndufb6, Ndufa4* and *Ndufb5*) and ATP synthase (*Atp5f1*)), protein turnover (Proteasome subunit alpha type 3-like (*Psma3l*)), oxidative stress response (Metallothionein-2 (*Mt2*)) and retinol metabolism (cellular retinol-binding protein (*Rbp7*)). This top-10 list indicates that FOS especially affects cellular energy metabolism in rat colonic mucosa, this was supported by the pathway analysis results, as described below.

Table 1. Top 10 of highest ranked genes by Random forest and by t-test.

Gene name	Sequence ID	Gene Symbol	Fold change*	p value
Glucagon gene, exon 6	K02813	*Glc*	2.6	5E-10
Cellular retinol-binding protein 7	P02696	*Crbp*	4.0	2E-09
NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 5	XM_215544	*Ndufb5*	1.3	3E-08
Unknown (LOC295337)	XM_215660		1.6	4E-08
NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 6	XM_216378	*Ndufb6*	1.5	7E-08
ATP synthase, H+ transporting, mitochondrial F0 complex, subunit b, isoform 1	NM_134365	*Atp5f1*	1.4	3E-07
Protein C11orf10, LOC309206	XM_219574	-	1.4	3E-07
NADH dehydrogenase (ubiquinone) 1 alpha subcomplex, 4	NM_010886	*Ndufa4*	1.4	3E-06
Metallothionein-2	BF556648	*Mt2*	1.9	4E-06
Proteasome subunit alpha type 3-like	BN000326	*Psma3l*	1.4	7E-06

*Fold change FOS/ control.

Pathway analysis

Pathway analysis identified processes most affected by dietary FOS in colonic mucosa. 366 Genes of the 997 selected genes could be classified based on GO term or based on GeneGo annotation in the Metacore database[24]. The most significant processes were an entire range of mitochondria related processes such as mitochondrial electron transport, oxidative phosphorylation, translation in mitochondria and proteins targeted to mitochondria (supplemental table 6.3). Another highly classified process was proteolysis (supplemental table 6.3). To prevent bias in biological interpretation due to gene selection (by t-test and RF), we also applied Gene Set Enrichment

Analysis (GSEA) which includes all genes in the dataset (28180 genes). We focused on curated gene-sets originating from GenMapp, Biocart, SigmaAldrich and Broad institute. Comparable biological processes were found by GSEA as observed in Metacore: again electron transport and oxidative phosphorylation were most significant, followed by proteasomal degradation (supplemental table 6.4). Thus the threshold based Metacore analysis and the threshold free enrichment analysis, GSEA, gave similar results for the most significantly changed processes by FOS. The results obtained by pathway programs consist of many overlapping pathways. We combined pathways with overlapping genes such as mitochondrial electron transport (Metacore), electron transport (Metacore) and electron transport chain (GSEA) and categorized the processes.

Analysis based on pathway programs is restricted to the well annotated genes[34]. As only 36% of the Agilent whole genome array is recognized by Metacore and only 35% by GSEA, we manually extended the significantly altered pathways with the non-recognized genes using literature and databases mining (using Biocarta, Source, Genecards). This strengthened the pathway outcome, as we were able to identify many additional genes affected by FOS that could be added to the processes already identified by the pathway programs. This was the case for transcription identified by Metacore as nucleosome assembly, cell turnover identified by GSEA as programmed cell death, cytoskeleton and vesicle related processes (muscle filament sliding and cytoskeleton-dependent intracellular transport in Metacore) and oxidative stress (free radical induced apoptosis in GSEA) (figure 2). In addition, we identified FOS affected genes that were not grouped into a pathway by both programs but obviously belong to the same biological process, this was the case for mucosal barrier, transport, and peptide hormones (figure 2).
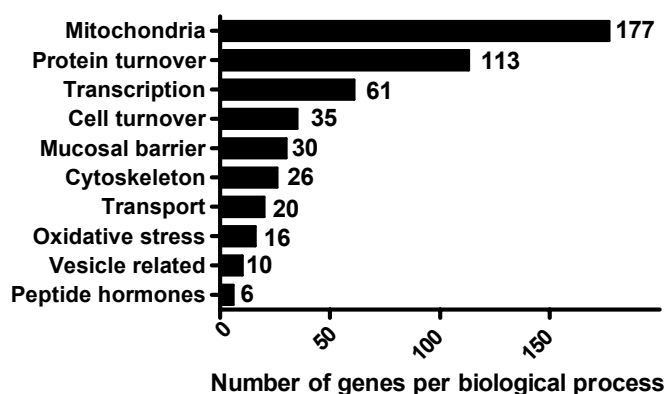
Figure 2. Classification of the genes affected by FOS into biological processes. Analyzed by Metacore, GSEA and data mining.

Detailed analysis of the mitochondrial processes showed an increased expression of genes associated with al five complexes of the OXPHOS complex, TCA-cycle and mitochondrial ribosomes and mitochondrial protein transport (table 2). In addition to protein degradation, which was found in the pathway programs, protein translation and maturation were also affected by FOS. From these 113 genes more than 90% showed increased expression upon FOS indicating increased protein turnover (table 2). Detailed gene expression data for all processes mentioned in table 2 is presented in Supplemental table 6.5 available at the website www.foodbioactives.nl.

Table 2. Detailed classification of biological processes affected by FOS.

Biological process		Number of genes affected by FOS*
Mitochondria	Complex I	27
	Complex II	5
	Complex III	2
	Complex IV	13
	Complex V	21
	Metabolism and TCA cycle	53
	Mitochondrial ribosomes	33
	Protein transport	11
	Miscellaneous	12
Protein turnover	Protein degradation	31
	Translation	44
	Protein maturation	32
	Miscellaneous	6
Transcription	Chromatin related	22
	mRNA metabolism	7
	Transcription	26
	Miscellaneous	6
Cell turnover	Apoptosis	19
	Growth/ differentiation	16
Mucosal barrier		30
Cytoskeleton		26
Transport		20
Oxidative stress		16
Vesicle related		10
Peptide hormones		6
Other†		503
		997

* genes with p-value<0.001 or selected by RF threshold.

† genes with less than 5 other genes belonging to the same process, unknown genes and, genes not part of a known process.

Confirmation of array results by Q-PCR

Confirmation of FOS induced processes was performed by Q-PCR. Genes from several FOS affected processes were analyzed by individual Q-PCR. We selected nine genes from mitochondria related processes: NADH dehydrogenase (ubiquinone) 1 beta subcomplex 9 (*Nduf9b*) , succinate dehydrogenase complex subunit B (*Sdhb*), ubiquinol-cytochrome c reductase binding protein (*UbiqcytC*), cytochrome c oxidase subunit VIIb *(Cox7b)*, ATP synthase H$^+$ transporting mitochondrial F0 complex subunit G *(ATP5g)*, aldo-keto reductase family 1 member B8 (*Akr1b8*), malic enzyme 1 (*Me1*), mitochondrial ribosomal protein S16 (*RiboS16*) , translocase of inner mitochondrial membrane 8 homolog b (*Timm8b*). In addition, one gene from protein degradation proteasome subunit alpha type 3-like (*Psma3l*), two genes related to mucosal barrier phospholipase A2, group IIA (platelets synovial fluid) (*Pla2g2a*) and trefoil factor 3 (*Tff3*) and one gene from the peptide hormones, proglucagon (*Gcg*) were analyzed. We additionally selected retinol binding protein 7 (*Rbp7*) as this gene was highest induced by FOS (4-fold). Q-PCR fully confirmed the microarray data (table 3). The p-value in the Q-PCR analysis reached significance (p<0.05) for 11 out of 13 genes. Two genes had a p-value>0.05 (*ATP 5g* and *Timm8b*).

Table 3. Q-PCR confirmation of microarray data.

Gene name	Gene symbol	sequence ID	Micro-array Ratio*	Q-PCR Ratio*	SEM±	p-value†
Mitochondria						
Complex I						
NADH dehydrogenase (ubiquinone) 1 beta subcomplex, 9	*Nduf9b*	XM_216929	1.47	1.29	0.05	0.001
Complex II						
Succinate dehydrogenase complex, subunit B, iron sulfur	*Sdhb*	XM_216558	1.43	1.32	0.05	<0.001
Comples III						
Ubiquinol-cytochrome c reductase binding protein	*Uqcrb*	XM_001074024	1.53	1.28	0.05	0.005
Comples IV						
Cytochrome c oxidase subunit VIIb	*Cox7b*	NM_182819	1.61	1.36	0.06	0.001
Comples V						
ATP synthase, H+ transporting, mitochondrial F0 complex, subunit G	*ATP5g*	XM_001075306	1.52	1.15	0.05	0.11
Metabolism and TCA cycle						
Aldo-keto reductase family 1, member B8	*Akr1b8*	NM_173136	2.13	2.15	0.22	0.006
malic enzyme 1	*Me1*		1.65	1.76	0.11	<0.001
Mitochondrial ribosomes						
Mitochondrial ribosomal protein S16	*Mrps16*	XM_001064095	1.43	1.20	0.05	0.03
Mitochondrial protein transport						
Translocase of inner mitochondrial membrane 8 homolog b	*Timm8b*	NM_022541	1.45	1.19	0.06	0.09
Protein turnover						
Proteasome subunit alpha type 3-like	*Psma31*	BN000326	1.39	1.24	0.04	0.002
Mucosal barrier						
Phospholipase A2, group IIA (platelets, synovial fluid)	*Pla2g2a*	NM_031598	3.73	4.70	0.87	0.03
Trefoil factor 3	*Tff3*	NM_013042	1.7	1.21	0.05	0.04
Top 10 gene						
Retinol binding protein 7	*Rbp7*	XM_575960	2.06	3.72	0.36	<0.001
Peptide hormone						
Proglucagon	*Gcg*	NM_012707	2.56	2.91	0.23	<0.001

* Ratio FOS diet/ Control diet.

± SEM of Q-PCR data is given (n= 12 per group).

† p-value of Q-PCR data is given, the p-value of microarray data were all <0.001.

Confirmation of gene expression results on protein level

To substantiate the FOS induced transcriptional modulation of mitochondrial genes at the protein level, we analyzed pooled mucosal scrapings of all rats (n=12 per group). The small increase in mRNA levels of complex IV subunits (~1.4 fold), was confirmed by a similar increase (1.5, 1.7 and 2.7 fold in independent pools, relative to actin) in protein levels of complex IV subunit COX II in the FOS group compared with the control group in pooled (n=12) mucosal scrapings (figure 3).
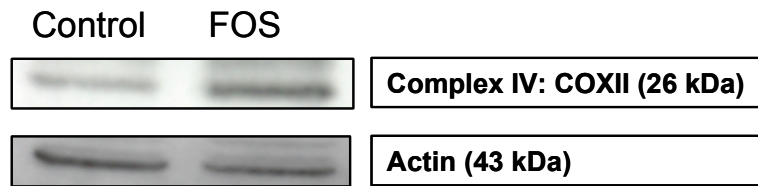
Control FOS	

	Complex IV: COXII (26 kDa)
	Actin (43 kDa)

Figure 3. Mucosal scapings (pool of n=12 per group) were examined for complex IV subunit COXII protein levels. The experiment was performed three times on separately prepared pools, showing a 1.5; 1.7 and 2.7 fold difference in COXII protein expression relative to actin, respectively. The 1.5 fold increase is shown.

The relatively high and significant induction of proglucagon gene by FOS was also examined at the protein level. Mucosal scrapings of randomly sampled control and FOS-fed rats (n=7 per group) were analyzed for GLP-1 protein levels by western blot and normalized to actin. FOS significantly increased GLP-1 protein levels in colonic mucosal tissue (Figure 4), substantiating the gene expression findings.

The clear effects of FOS on mitochondrial processes was not expected beforehand, therefore no precautions were taken at the time of sampling and storage that would allow post-hoc analysis of ATP levels. Alternatively, we attempted to examine levels of phosphorylated AMP-activated protein kinase (AMPK) which reflects the ADP/ATP ratio in cells[30]. Low levels of total AMPK protein could be detected in colonic scrapings with no difference between control and FOS (data not shown). AMPK phosphorylation could not be detected in mucosal scrapings, most probably due to the instable phosphorylations that require specific handling during the collection of the samples.
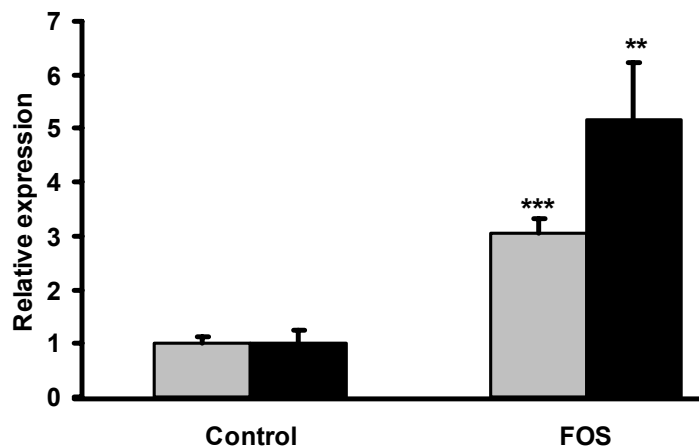
Figure 4. Relative expression of proglucagon mRNA (gray bar) and GLP-1 protein (black bar) in colonic mucosa of a random selection of control fed (n=7) and FOS fed (n=7) rats. mRNA and protein levels were normalized to actin levels. (**p<0.01, ***p<0.001).

Discussion

The CrEDTA results showed that dietary FOS increased intestinal permeability in rats in accordance with our previous study[77]. The FOS induced increase in intestinal permeability reduces intestinal barrier function as reported earlier[77]. Individual gene expression of 12 rats after ≈2 weeks FOS versus control feeding were explored on whole genome level and showed that the increased permeability could not be explained by changes in genes belonging to the tight junction system. No significant changes were observed in claudin 2 and 4, cadherins or tight junction protein 1. With our focus on gene expression, possible changes in protein levels and cellular

localization or modification cannot be excluded. 19 Genes related to apoptosis were affected by FOS with only modest fold changes. Although some pro-apoptotic genes were mildly affected (eg *Bax,* Caspase 7), no changes were seen in many key pro-apoptotic genes such as *Bad*, Caspase 3 and Apaf1. Therefore, we feel that apoptosis is not the main cause of the increased intestinal permeability observed. Known mucosal defense genes such as defensins, mucins and calprotectin were also unaffected by FOS. These few and small transcriptional changes in potential barrier related genes cannot explain the profound and consistent effects of FOS on intestinal permeability in rats.

On the other hand, we identified multiple genes associated with energy metabolism (177 mitochondria related genes) that were significantly modified by FOS. Protein turnover was also clearly affected by FOS (113 genes). Coincidence of increased permeability and upregulation of these genes suggests that these processes play a major role in preservation of intestinal mucosal integrity.

The most striking observation was the induction of a broad range of mitochondrial genes. Increased expression of mitochondrial genes is commonly observed during disturbed ATP homeostasis caused by increased energy demand or decreased mitochondrial energy supply[84]. ATP depletion by a wide range of uncoupling agents induces mitochondrial genes such as cyclo-oxidase, NADH dehydrogenase and 16S mitochondrial mRNA in rat colon[58], cytochrome c oxidase subunit III (*CoxIII*) in rat heart[84], Cytochrome c, *CoxII* and Atp-ase in fibroblasts and several mouse tissues and *CoxIV* and Adenine nucleotide translocator (*Ant)* and liver cell culture studies [19,62]. FOS induced most of the above mentioned genes (*Ant*, 24 NADH dehydrogenases, 7 *Cox* subunits and 3 ATP-ases). This could be confirmed at protein level for complex IV subunit II. Together this strongly indicates that FOS caused ATP depletion in colonic epithelial cells.

The next question is whether the increased expression of mitochondrial genes by FOS, which probably reflects ATP-depletion, is responsible for the observed increase in intestinal permeability in FOS-fed rats. Many studies using epithelial cell lines have shown that ATP depletion is a cause of paracellular hyperpermeability[42,43,79]. Another strong indication that mitochondria are important in maintaining intestinal permeability is derived from studies on chronic non-steroidal anti-inflammatory drugs (NSAIDs)[69]. Direct exposure of rat intestine to the NSAID indomethacin or the mitochondrial uncoupler 2,4-dinitrophenol (DNP) increases epithelial permeability [51]. Also, NSAIDs or DNP induced uncoupling of intestinal mitochondria was shown to lead to increased bacterial translocation in rats and intestinal cell lines, and immune cell infiltration and ulceration in rats[51,70]. Enhanced bacterial translocation reflects impaired barrier function. The NSAID induced increased permeability is attenuated with co-administration of glucose and citrate, substrates for tricarboxylic acid cycle and glycolysis, or ATP[8,10].

What could be the mechanism of the FOS induced increase in intestinal mitochondrial gene expression and possible ATP-depletion, and the increased mucosal permeability? Previous studies in rats[37,60,76] and humans[75] showed rapid fermentation of FOS in cecum and colon by the endogenous microflora resulting in lactic acid accumulation, SCFA production, and decreased pH of luminal contents. Although luminal production of modest quantities of SCFAs is essential for normal colonic mucosal function[64], overproduction or accumulation of SCFA along with low pH

in the intestinal lumen has been shown to cause intestinal injury leading to increased intestinal permeability[3,39,48]. In vitro studies show that increased permeability of intestinal epithelial monolayers caused by exposure to excess SCFA levels and reduced extracellular pH is associated with cellular ATP depletion[36,45]. The SCFA induced depletion of cellular ATP coincided with a reduced intracellular pH (pH_i) in perfused livers[5]. In an acidic luminal environment, a relatively larger portion of SCFA becomes protonated, facilitating passive diffusion of SCFA across the apical hydrophobic enterocyte membrane causing intracellular acidification[13,15]. The above mentioned studies were done in vitro or in perfused liver, but the concentrations applied can impair the pH_i homeostasis in colonocytes in FOS fed rats[37,47]. Taken together, FOS reduces luminal pH and increases levels of SCFAs, this likely leads to acidification of the cellular cytoplasm (decrease in pH_i) and subsequent ATP-depletion. ATP-depletion can occur because the decrease in pH_i is compensated by H^+ excretion in exchange for Na^+ by the pH_i regulating $Na^+/HCO3^-$ cotransporter (NBC) and the Na^+/H^+ exchanger (NHE) in the basolateral membrane [18,23]. This indirectly activates the Na+,K+-ATPase (ATP1) which is known to require ~25% of the cellular ATP turnover under basal conditions[11]. Long-term exposure of enterocytes to high SCFA concentrations under low pH conditions might therefore disturb or exhaust plasma membrane pumps, leading to rapid ATP-depletion[36](BOX 1). FOS diet significantly increased expression of *Atp1b1* (1.2-fold, p<0.001). No changes were found in the gene expression of the transporters *Nbc* and *Nhe* possible increased activity of these transporters can occur without concomitant changes in mRNA gene expression. A strong indication supporting SCFA and low luminal pH as inducers of the observed effects, are results of previous studies of our lab showing that the adverse affects of FOS on mucosal barrier, i.e. increased permeability and decreased resistance to pathogens, are absent when calcium is supplemented to the diets[74]. Calcium is known to increase the buffering capacity of luminal contents by its precipitation with dietary phosphate. By preventing acidification during fermentation, SCFA will remain in the anionic form, for which the cellular plasma membrane is not permeable.

Another possible, but less likely, mechanism of FOS induced mitochondrial gene expression is SCFA induced accumulation of Acyl-CoA in mitochondria. This can result in depletion of ATP in the intramitochondrial compartment and accumulation of AMP[67]. We observed induction of Acyl-Coa synthetase (FC1.3, p<0.001). Furthermore, SCFAs are reported to induce opening of mitochondrial inner membrane channels like the permeability transition pore complex resulting in loss of membrane potential and consequently ATP depletion[32,36]. FOS significantly changed the expression of pore complex member adenine nucleotide translocator (*Ant*) (FC 1.3, p<0.001). However no significant changes were observed in the expression of other members such as voltage activated anion channel, Cyclophilin D and peripheral benzodiazipine receptor.

Although the exact mechanism is not clear, the observed increase of many mitochondria related genes in the present study seem to represent induced mitochondrial activity to compensate for SCFA -induced ATP depletion. This is supported by the fact that all five OXPHOS complexes, mitochondrial ribosomes and many other mitochondria related genes were induced, indicating an overall increase in mitochondrial capacity. We would like to add that although in vitro studies and studies performed in other tissues show ATP depletion due to SCFA exposure, this has to be confirmed for the colon in the intact animal, preferably by direct assessment of ATP levels.

BOX 1. Proposed mechanism of dietary FOS induced intestinal permeability
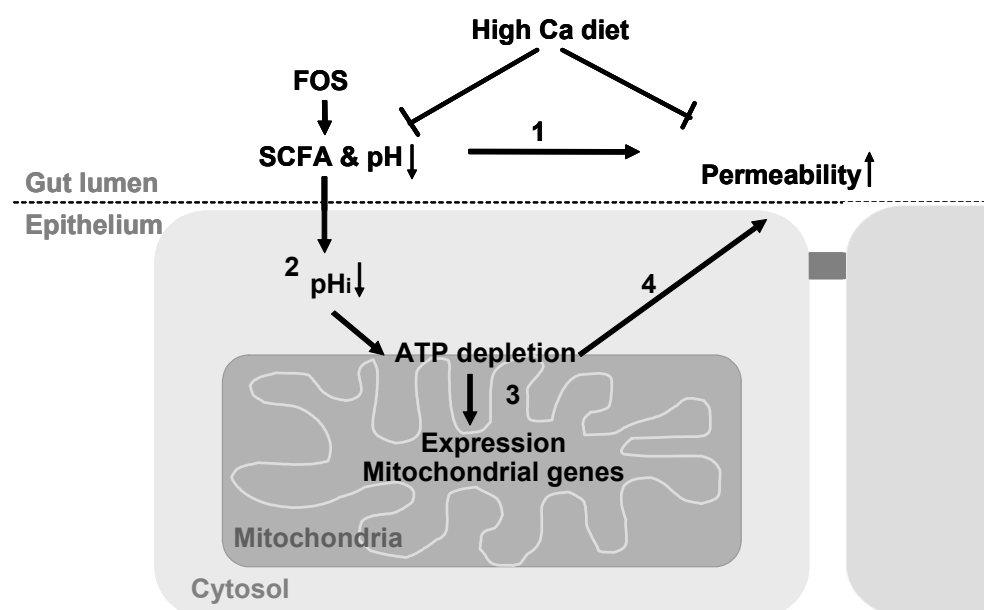
1 High levels of FOS fermentation products increase intestinal permeability in vivo [3,39,77].

2 Excess SCFAs cause intracellular acidification of epithelial cells. When protonated-SCFA diffuse from the gut lumen into epithelial cells [15,29]. The SCFA cause intracellular acidification and induce proton pump activity (NHE and NBC transporters) which may lead to ATP depletion [5,36]

3 Reduced ATP levels, by increased energy demand, chronic mitochondrial uncoupling or any other cause of disturbed energy metabolism, are compensated by increased mitochondrial gene expression and mitochondrial biogenesis [62,84].

4 Disturbed energy metabolism leads to increased permeability. In agreement: ATP-depletion in epithelial cell lines causes paracellular hyperpermeability [42,43,79] and uncoupling of intestinal mitochondria leads to increased bacterial translocation, immune cell infiltration and ulceration in rats [51,69].

Calcium supplementation of a FOS diet counteracts FOS induced intestinal permeability. Calcium prevents acidification of intestinal contents during fermentation and thus formation of protonated-SCFA.

The second major process affected by FOS was protein turnover. FOS induced 27 ribosomal proteins and 30 proteasomeal genes, indicating increased protein turnover. Intracellular proteins are targeted to the proteasomal degradation system by ubiquitination[50]. FOS induced 9 ubiquitin related genes, including several isoforms of the E2 ubiquitin carrier enzyme. Proteasomal degradation is tightly controlled and removes denatured, misfolded and damaged proteins. The clear increase in proteasomal gene expression might result from increased presence of misfolded proteins. One common cause of misfolding of cellular proteins is mild oxidative stress[33]. FOS induced several genes related to oxidative stress, such as metallothionein-2 (*Mt2*) metallothionein-1a (*Mt1a*), six glutathione S-transferases (*Gst*'s), heme oxygenase 1 (*Hmox1*), and superoxide dismutase 1 and 3 (*Sod1* and *Sod3*). *Mt2* is one of the top-10 most regulated genes by FOS, it protects against oxidative stress by capturing harmful oxygen radicals by its cystein residues[78]. An increase in oxidative stress proteins is an indirect marker for production of reactive oxygen species (ROS)[65]. It is well recognized that induced mitochondrial activity can increase production of ROS. The increased expression of oxidative stress genes and mitochondrial genes observed in the present study suggests increased mitochondrial activity, possibly associated with increased ROS production and increased in protein oxidation. Oxidized proteins are often misfolded, and directed to proteasomes for degradation. In support, the translocase *Sec61* responsible for

intracellular transport of misfolded proteins from the ER to the proteasome was significantly increased 1.4 fold (p<0.001) by FOS.

Increased proteasomal gene expression might also reflect the formation of immunoproteasomes (I-proteasomes). The I-proteasome plays a role in antigen processing and is composed of the 20S proteasome complex coupled to 2 homologous complexes called PA28-α (*Psme1*) and -β *Psme2*[81]. FOS supplementation significantly increased both *Psme* subunits. I-proteasome derived proteolysed antigen fragments are loaded on class I MHC and presented to receptors on CD8[+] T-cells leading to activation of an immune response via NF-κb[81].The expression of PA28 is known to be induced by cytokines like IFNγ and microbial infection[71]. Alterations in the I-proteasome in injured intestinal epithelium are observed in colon biopsies of IBD patients and IBD-mouse models[26,81]. The increased intestinal permeability due to the FOS diet can cause increased exposure of the mucosa to bacteria and therefore induce I-proteasome gene expression.

The top 10 most significantly changed genes by FOS showed to be good representatives of the major biological processes selected from pathway analysis. Four of these ten are members of mitochondria related processes (*Ndufb6*, NADH-ubiquinone oxidoreductase *Mlrq, Ndufb5 and Atp5f1*). Another top 10 member, *Psma3l,* represented the second biological process induced by FOS, namely protein turnover. In addition to these genes, the gene coding for cellular retinol-binding protein (*Rbp7*) was the most highly induced gene (4-fold) and another cellular retinol binding protein, *Rbp2* was induced 2-fold by FOS. RBPs are required for uptake, intracellular transport and metabolism of vitamin A. Vitamin A is a fat-soluble vitamin necessary for growth and differentiation of epithelial tissues. RBP7 and RBP2 belong to the fatty-acid binding protein (FABP) family. FOS also significantly increased *Fabp1* 2.9 fold. At present, we cannot explain the FOS induced expression of these genes and its relation to the functional effects observed in this and our earlier FOS studies.

Proglucagon is one of the most highly induced gene by FOS (2.6-fold, p<0.001). The proglucagon gene is a precursor encoding several glucagon-like peptides. In intestinal enteroendocrine cells the gene codes for oxyntomodulin, GLP1 and GLP2[31]. This gene was previously found to be induced by non digestible oligosaccharides[12]. Our study confirms that dietary FOS increases proglucagon gene expression and GLP-1 protein expression in the colon mucosa of rats. Besides FOS induced expression of proglucagon, FOS induced the expression of several other gut-derived peptide hormones, namely cholecystokinin (*Cck*), peptide YY (*Pyy*) and pancreatic polypeptide (*Ppy*). Increase of PYY has previously been reported in rat colon by SCFA[86]. Proglucagon, PYY, PPY and CCK are all expressed by enteroendocrine L cells in colon and play a role in gut-nutrient sensing[16]. In the hypothalamus nutrient sensing is also regulated by these hormones and directly related to ATP status[38]. It is speculated that the same mechanism is applicable to enteroendocrine L cells in de gut[86]. Since in our study, increased gene expression of these 4 peptide hormone genes coincides with alterations in mitochondrial processes, it is tempting to speculate that these hormones also influence or respond to energy metabolism in intestinal epithelial cells. Besides a role in energy homeostasis, GLP2, PYY and CCK have growth-promoting properties on the intestinal epithelium in vivo[25]. GLP1 stimulated cell proliferation has been reported for liver and pancreas[21]. GLP2 is involved in regulation of mucosal epithelial integrity[22]. It stimulates intestinal crypt cell proliferation[27] and reduces

apoptosis, therefore enhances mucosal regeneration. It has beneficial effects on many causes of intestinal injury, such as stress, vascular ischemia, NSAID administration and chemically induced injury in rodents, and decreases subsequent intestinal permeability[35,55]. The upregulation of these genes gene might thus be a response to the impaired intestinal barrier in FOS-fed rats.

FOS consistently increased intestinal permeability, but the present study showed hardly any effect on expression of well known intestinal integrity genes. Most surprisingly no changes were observed in genes related to tight junctions that were expected since tight junctions are key regulators of paracellular transport. However, changes in epithelial permeability are a result of internalization of the tight-junction proteins occludin, claudin and junctional adhesion molecule-A[66]. These cellular translocations can occur without concomitant changes in mRNA gene expression. Detection of such effects would require a different approach from transcriptomics. Imunohistochemistry could show whether translocation of TJ proteins occurred in the FOS fed rats compared with control fed rats and is under current investigation.

Altogether we show that altered barrier integrity induced by FOS-diet coincides with a clear increase in mitochondrial gene expression, suggesting that mitochondrial energy metabolism is important for maintaining the intestinal barrier. The role of mitochondria in maintenance of the intestinal barrier is already accepted in NSAID or DNP uncoupling studies. We speculate that excess production of SCFA and acidification of luminal contents results in SCFA induced ATP depletion of colonic epithelial cells. Insight into the role of mitochondrial function and ATP depletion is of relevance, not only for the application of FOS and other prebiotics in food products on the current market, but especially for mechanistic understanding of intestinal disorders where gut permeability changes are observed.

Acknowledgments

References

1.   Allison, D.B., Cui, X., Page, G.P. & Sabripour, M. Microarray data analysis: from disarray to consolidation and consensus. *Nat Rev Genet* 7, 55-65 (2006).
2.   Andersson, H.B., Ellegard, L.H. & Bosaeus, I.G. Nondigestibility characteristics of inulin and oligofructose in humans. *J Nutr* 129, 1428S-1430S (1999).
3.   Argenzio, R.A. & Meuten, D.J. Short-chain fatty acids induce reversible injury of porcine colon. *Dig Dis Sci* 36, 1459-1468 (1991).
4.   Arslan, G., Atasever, T., Cindoruk, M. & Yildirim, I.S. (51)CrEDTA colonic permeability and therapy response in patients with ulcerative colitis. *Nucl Med Commun* 22, 997-1001 (2001).

5.   Beauvieux, M.C., Tissier, P., Gin, H., Canioni, P. & Gallis, J.L. Butyrate impairs energy metabolism in isolated perfused liver of fed rats. *J Nutr* 131, 1986-1992 (2001).

6.   Benjamini, Y.H., Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc.* 57, 289-300 (1995).

7.   Bhatia, V. & Tandon, R.K. Stress and the gastrointestinal tract. *J Gastroenterol Hepatol* 20, 332-339 (2005).

8.   Bjarnason, I., *et al.* Glucose and citrate reduce the permeability changes caused by indomethacin in humans. *Gastroenterology* 102, 1546-1550 (1992).

9.   Bouhnik, Y., *et al.* Administration of transgalacto-oligosaccharides increases fecal bifidobacteria and modifies colonic fermentation metabolism in healthy humans. *J Nutr* 127, 444-448 (1997).

10.   Bours, M.J., Troost, F.J., Brummer, R.J., Bast, A. & Dagnelie, P.C. Local effect of adenosine 5'-triphosphate on indomethacin-induced permeability changes in the human small intestine. *Eur J Gastroenterol Hepatol* 19, 245-250 (2007).

11.   Buck, L.T. & Hochachka, P.W. Anoxic suppression of Na(+)-K(+)-ATPase and constant membrane potential in hepatocytes: support for channel arrest. *Am J Physiol* 265, R1020-1025 (1993).

12.   Cani, P.D., Hoste, S., Guiot, Y. & Delzenne, N.M. Dietary non-digestible carbohydrates promote L-cell differentiation in the proximal colon of rats. *Br J Nutr* 98, 32-37 (2007).

13.   Charney, A.N., Micic, L. & Egnor, R.W. Nonionic diffusion of short-chain fatty acids across rat colon. *Am J Physiol* 274, G518-524 (1998).

14.   Chen, J.J., Wang, S.J., Tsai, C.A. & Lin, C.J. Selection of differentially expressed genes in microarray data analysis. *Pharmacogenomics J* (2006).

15.   Chu, S. & Montrose, M.H. Transepithelial SCFA fluxes link intracellular and extracellular pH regulation of mouse colonocytes. *Comp Biochem Physiol A Physiol* 118, 403-405 (1997).

16.   Cox, H.M. Peptide YY: a neuroendocrine neighbor of note. *Peptides* 28, 345-351 (2007).

17.   Cran-website. http://cran.r=project.org/.

18.   DeSoignie, R. & Sellin, J.H. Propionate-initiated changes in intracellular pH in rabbit colonocytes. *Gastroenterology* 107, 347-356 (1994).

19.   Desquiret, V., *et al.* Dinitrophenol-induced mitochondrial uncoupling in vivo triggers respiratory adaptation in HepG2 cells. *Biochim Biophys Acta* 1757, 21-30 (2006).

20.   Diaz-Uriarte, R. & Alvarez de Andres, S. Gene selection and classification of microarray data using random forest. *BMC Bioinformatics* 7, 3 (2006).

21.   Doyle, M.E. & Egan, J.M. Mechanisms of action of glucagon-like peptide 1 in the pancreas. *Pharmacol Ther* 113, 546-593 (2007).

22.   Dube, P.E. & Brubaker, P.L. Frontiers in glucagon-like peptide-2: multiple actions, multiple mediators. *Am J Physiol Endocrinol Metab* 293, E460-465 (2007).

23.   Durand, T., Gallis, J.L., Masson, S., Cozzone, P.J. & Canioni, P. pH regulation in perfused rat liver: respective role of Na(+)-H+ exchanger and Na(+)-HCO3- cotransport. *Am J Physiol* 265, G43-50 (1993).

24.   Ekins, S., Nikolsky, Y., Bugrim, A., Kirillov, E. & Nikolskaya, T. Pathway mapping tools for analysis of high content data. *Methods Mol Biol* 356, 319-350 (2007).

25.   Estall, J.L. & Drucker, D.J. Tales beyond the crypt: glucagon-like peptide-2 and cytoprotection in the intestinal mucosa. *Endocrinology* 146, 19-21 (2005).

26.   Fitzpatrick, L.R., Khare, V., Small, J.S. & Koltun, W.A. Dextran sulfate sodium-induced colitis is associated with enhanced low molecular mass polypeptide 2 (LMP2) expression and is attenuated in LMP2 knockout mice. *Dig Dis Sci* 51, 1269-1276 (2006).

27.   Ghatei, M.A., *et al.* Proglucagon-derived peptides in intestinal epithelial proliferation: glucagon-like peptide-2 is a major mediator of intestinal epithelial proliferation in rats. *Dig Dis Sci* 46, 1255-1263 (2001).

28.   Gibson, G.R., Beatty, E.R., Wang, X. & Cummings, J.H. Selective stimulation of bifidobacteria in the human colon by oligofructose and inulin. *Gastroenterology* 108, 975-982 (1995).

29.   Gonda, T., Maouyo, D., Rees, S.E. & Montrose, M.H. Regulation of intracellular pH gradients by identified Na/H exchanger isoforms and a short-chain fatty acid. *Am J Physiol* 276, G259-270 (1999).

30.   Hardie, D.G. The AMP-activated protein kinase pathway--new players upstream and downstream. *J Cell Sci* 117, 5479-5487 (2004).

31.   Holst, J.J. Enteroglucagon. *Annu Rev Physiol* 59, 257-271 (1997).

32.   Jan, G., *et al.* Propionibacteria induce apoptosis of colorectal carcinoma cells via short-chain fatty acids acting on mitochondria. *Cell Death Differ* 9, 179-188 (2002).

33.   Jung, T., Bader, N. & Grune, T. Oxidized proteins: intracellular distribution and recognition by the proteasome. *Arch Biochem Biophys* 462, 231-237 (2007).

34.   Khatri, P. & Draghici, S. Ontological analysis of gene expression data: current tools, limitations, and open problems. *Bioinformatics* 21, 3587-3595 (2005).

35.   Kouris, G.J.*, et al.* The effect of glucagon-like peptide 2 on intestinal permeability and bacterial translocation in acute necrotizing pancreatitis. *Am J Surg* 181, 571-575 (2001).

36.   Lan, A., Lagadic-Gossmann, D., Lemaire, C., Brenner, C. & Jan, G. Acidic extracellular pH shifts colorectal cancer cell death from apoptosis to necrosis upon exposure to propionate and acetate, major end-products of the human probiotic propionibacteria. *Apoptosis* 12, 573-591 (2007).

37.   Le Blay, G.M., Michel, C.D., Blottiere, H.M. & Cherbut, C.J. Raw potato starch and short-chain fructo-oligosaccharides affect the composition and metabolic activity of rat intestinal microbiota differently depending on the caecocolonic segment involved. *J Appl Microbiol* 94, 312-320 (2003).

38.   Lee, K., Li, B., Xi, X., Suh, Y. & Martin, R.J. Role of neuronal energy status in the regulation of adenosine 5'-monophosphate-activated protein kinase, orexigenic neuropeptides expression, and feeding behavior. *Endocrinology* 146, 3-10 (2005).

39.   Lin, J.*, et al.* Variable effects of short chain fatty acids and lactic acid in inducing intestinal mucosal injury in newborn rats. *J Pediatr Gastroenterol Nutr* 35, 545-550 (2002).

40.   Lunetta, K.L., Hayward, L.B., Segal, J. & Van Eerdeweegh, P. Screening large-scale association study data: exploiting interactions using random forests. *BMC Genet* 5, 32 (2004).

41.   Lyons-Weiler, J.*, et al.* Assessing the statistical significance of the achieved classification error of classifiers constructed using serum peptide profiles, and a prescription for random sampling repeated studies for massive high-throughput genomic and proteomics studies. *Cancer Informatics* 1, 53-77 (2005).

42.   Madsen, K.L., Yanchar, N.L., Sigalet, D.L., Reigel, T. & Fedorak, R.N. FK506 increases permeability in rat intestine by inhibiting mitochondrial function. *Gastroenterology* 109, 107-114 (1995).

43.   Mandel, L.J., Bacallao, R. & Zampighi, G. Uncoupling of the molecular 'fence' and paracellular 'gate' functions in epithelial tight junctions. *Nature* 361, 552-555 (1993).

44.   Mankertz, J. & Schulzke, J.D. Altered permeability in inflammatory bowel disease: pathophysiology and clinical implications. *Curr Opin Gastroenterol* 23, 379-383 (2007).

45.   Menconi, M.J.*, et al.* Acidosis induces hyperpermeability in Caco-2BBe cultured intestinal epithelial monolayers. *Am J Physiol* 272, G1007-1021 (1997).

46.   Montalto, M.*, et al.* Lactobacillus acidophilus protects tight junctions from aspirin damage in HT-29 cells. *Digestion* 69, 225-228 (2004).

47.   Moreau, N.M.*, et al.* Restoration of the integrity of rat caeco-colonic mucosa by resistant starch, but not by fructo-oligosaccharides, in dextran sulfate sodium-induced experimental colitis. *Br J Nutr* 90, 75-85 (2003).

48.   Nafday, S.M.*, et al.* Short-chain fatty acids induce colonic mucosal injury in rats with various postnatal ages. *Pediatr Res* 57, 201-204 (2005).

49.   Nagura, H. Mucosal defense mechanism in health and disease. Role of the mucosal immune system. *Acta Pathol Jpn* 42, 387-400 (1992).

50.   Nandi, D., Tahiliani, P., Kumar, A. & Chandu, D. The ubiquitin-proteasome system. *J Biosci* 31, 137-155 (2006).

51.   Nazli, A.*, et al.* Epithelia under metabolic stress perceive commensal bacteria as a threat. *Am J Pathol* 164, 947-957 (2004).

52.   NCBI GEO website. http://www.ncbi.nlm.nih.gov/geo/. (,).

53.   Pellis, L., Franssen-van Hal, N.L., Burema, J. & Keijer, J. The intraclass correlation coefficient applied for evaluation of data correction, labeling methods, and rectal biopsy sampling in DNA microarray experiments. *Physiol Genomics* 16, 99-106 (2003).

54.   Potten, C.S. & Booth, C. The role of radiation-induced and spontaneous apoptosis in the homeostasis of the gastrointestinal epithelium: a brief review. *Comp Biochem Physiol B Biochem Mol Biol* 118, 473-478 (1997).

55.   Prasad, R., Alavi, K. & Schwartz, M.Z. Glucagonlike peptide-2 analogue enhances intestinal mucosal mass after ischemia and reperfusion. *J Pediatr Surg* 35, 357-359 (2000).

56.   Quackenbush, J. Extracting meaning from functional genomics experiments. *Toxicol Appl Pharmacol* 207, 195-199 (2005).

57.   R development core team. R: A language and environment for statistical computing. in *http://www.R-project.org* (2004).

58.   Rachamim, N.*, et al.* Dexamethasone enhances expression of mitochondrial oxidative phosphorylation genes in rat distal colon. *Am J Physiol* 269, C1305-1310 (1995).

59.   Reeves, P.G., Nielsen, F.H. & Fahey, G.C., Jr. AIN-93 purified diets for laboratory rodents: final report of the American Institute of Nutrition ad hoc writing committee on the reformulation of the AIN-76A rodent diet. *J Nutr* 123, 1939-1951 (1993).

60. Remesy, C., Levrat, M.A., Gamet, L. & Demigne, C. Cecal fermentations in rats fed oligosaccharides (inulin) are modulated by dietary calcium level. *Am J Physiol* 264, G855-862 (1993).

61. Rodenburg, W.*, et al.* Salmonella induces prominent gene expression in the rat colon. *BMC Microbiol* 7, 84 (2007).

62. Rohas, L.M.*, et al.* A fundamental system of cellular energy homeostasis regulated by PGC-1alpha. *Proc Natl Acad Sci U S A* 104, 7933-7938 (2007).

63. Rubin, E. Circumventing the cut-off for enrichment analysis. *Brief Bioinform* 7, 202-203 (2006).

64. Scheppach, W. Effects of short chain fatty acids on gut morphology and function. *Gut* 35, S35-38 (1994).

65. Scherz-Shouval, R. & Elazar, Z. ROS, mitochondria and the regulation of autophagy. *Trends Cell Biol* (2007).

66. Schneeberger, E.E. & Lynch, R.D. The tight junction: a multifunctional complex. *Am J Physiol Cell Physiol* 286, C1213-1228 (2004).

67. Schonfeld, P., Wojtczak, A.B., Geelen, M.J., Kunz, W. & Wojtczak, L. On the mechanism of the so-called uncoupling effect of medium- and short-chain fatty acids. *Biochim Biophys Acta* 936, 280-288 (1988).

68. Schulzke, J.D.*, et al.* Disrupted barrier function through epithelial cell apoptosis. *Ann N Y Acad Sci* 1072, 288-299 (2006).

69. Somasundaram, S.*, et al.* Mitochondrial damage: a possible mechanism of the "topical" phase of NSAID induced injury to the rat intestine. *Gut* 41, 344-353 (1997).

70. Somasundaram, S.*, et al.* Uncoupling of intestinal mitochondrial oxidative phosphorylation and inhibition of cyclooxygenase are required for the development of NSAID-enteropathy in the rat. *Aliment Pharmacol Ther* 14, 639-650 (2000).

71. Strehl, B.*, et al.* Interferon-gamma, the functional plasticity of the ubiquitin-proteasome system, and MHC class I antigen processing. *Immunol Rev* 207, 19-30 (2005).

72. Subramanian, A.*, et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102, 15545-15550 (2005).

73. Swidsinski, A.*, et al.* Viscosity gradient within the mucus layer determines the mucosal barrier function and the spatial organization of the intestinal microbiota. *Inflamm Bowel Dis* 13, 963-970 (2007).

74. Ten Bruggencate, S.J., Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L., Katan, M.B. & Van Der Meer, R. Dietary fructo-oligosaccharides and inulin decrease resistance of rats to salmonella: protective role of calcium. *Gut* 53, 530-535 (2004).

75. Ten Bruggencate, S.J., Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L., Katan, M.B. & van der Meer, R. Dietary fructooligosaccharides affect intestinal barrier function in healthy men. *J Nutr* 136, 70-74 (2006).

76. Ten Bruggencate, S.J., Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L. & Van der Meer, R. Dietary fructo-oligosaccharides dose-dependently increase translocation of salmonella in rats. *J Nutr* 133, 2313-2318 (2003).

77. Ten Bruggencate, S.J., Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L. & Van der Meer, R. Dietary fructooligosaccharides increase intestinal permeability in rats. *J Nutr* 135, 837-842 (2005).

78. Thirumoorthy, N., Manisenthil Kumar, K.T., Shyam Sundar, A., Panayappan, L. & Chatterjee, M. Metallothionein: an overview. *World J Gastroenterol* 13, 993-996 (2007).

79. Unno, N.*, et al.* Hyperpermeability and ATP depletion induced by chronic hypoxia or glycolytic inhibition in Caco-2BBe monolayers. *Am J Physiol* 270, G1010-1021 (1996).

80. van Hal, N.L.*, et al.* The application of DNA microarrays in gene expression analysis. *J Biotechnol* 78, 271-280 (2000).

81. Wang, J. & Maldonado, M.A. The ubiquitin-proteasome system and its role in inflammatory and autoimmune diseases. *Cell Mol Immunol* 3, 255-261 (2006).

82. Wang, Q., Pantzar, N., Jeppsson, B., Westrom, B.R. & Karlsson, B.W. Increased intestinal marker absorption due to regional permeability changes and decreased intestinal transit during sepsis in the rat. *Scand J Gastroenterol* 29, 1001-1008 (1994).

83. Wehkamp, J.*, et al.* Paneth cell antimicrobial peptides: topographical distribution and quantification in human gastrointestinal tissues. *FEBS Lett* 580, 5344-5350 (2006).

84. Wiesner, R.J.*, et al.* Stimulation of mitochondrial gene expression and proliferation of mitochondria following impairment of cellular energy transfer by inhibition of the phosphocreatine circuit in rat hearts. *J Bioenerg Biomembr* 31, 559-567 (1999).

85. Zeissig, S.*, et al.* Changes in expression and distribution of claudin 2, 5 and 8 lead to discontinuous tight junctions and barrier dysfunction in active Crohn's disease. *Gut* 56, 61-72 (2007).

86. Zhou, J.*, et al.* Peptide YY and proglucagon mRNA expression patterns and regulation in the gut. *Obesity (Silver Spring)* 14, 683-689 (2006).

Chapter 7

General discussion

General discussion

1 Introduction

One of the current challenges in nutritional science is to improve gut health by dietary modulation. To improve gut health, insight into the molecular mechanisms of the gut mucosal barrier is required, since disturbance of these molecular barrier mechanisms leads to disease. Increased knowledge of gut mucosal barrier mechanisms can, for example, aid in understanding intestinal disease processes, which is necessary for the development of functional foods or medications that promote gut health. Moreover, in-depth knowledge of mechanisms can provide biomarkers that enable monitoring of gut health in humans.

The importance of the gut to health and disease is reflected in the large amount of research on the gut. Most of this research has focused on the intestinal barrier in relation to diseases such as chronic inflammatory bowel disease and infectious disease[26] or in relation to nutritional modulation, for example using probiotics and prebiotics[18]. Despite this large amount of attention, key molecular mechanisms that determine a healthy barrier are not completely understood due to the complexity of the barrier. The barrier is complex because it is composed of several components that constantly interact with each other, such as the epithelial cells, the mucus layer, the immune system and the microflora. However, most research on the barrier has concentrated on only one or two of these components. Few studies have focused on the intact gut mucosal barrier in a living organism. This last approach is required to really understand the gut mucosal barrier. We therefore chose to study the complete barrier complex in a living organism.

Examining the function and key mechanisms of a healthy biological cell, tissue or organ generally requires looking for differences between the healthy control condition and a treatment or disease condition. Mechanisms in the intestine that are activated by treatment or disease are important for barrier system homeostasis. For the intestine this means that the barrier must be stimulated with different stress challenges to test its capacity to withstand those challenges; this reveals molecular responses that are important for the gut mucosal barrier. Examples of stress challenges for the intestinal barrier are infection with pathogens like Salmonella and exposure to nutrients like FOS. Previous functional observations have shown that infection results in translocation of Salmonella across the intestinal barrier to extra-intestinal organs and that dietary FOS induces increased intestinal permeability[76,77].

Salmonella is a frequently used model pathogen in gut barrier research[53,68]. One reason for this is that Salmonella is a main cause of human foodborne infectious disease. Despite the large amount of data on host molecular responses to Salmonella infection, a full overview of the *in vivo* molecular mucosal response is lacking, since most data are derived from *in vitro* cell culture studies or *ex vivo* model systems[33].

Studies on the effects of FOS on the intestine clearly show effects on the intestinal barrier of rats. For example, it increases both Salmonella translocation and intestinal permeability[76,77]. The mechanisms responsible for these effects of FOS on the intestine are unknown, but are probably related to enhanced colonic fermentation and production of organic acids by the gut microflora. Because of their clear effects on the intestinal barrier, both Salmonella and FOS are relevant gut

barrier challenges and are therefore useful as tools or models to examine processes important for barrier functioning.

Since little is known about the various intestinal mechanisms that are influenced by Salmonella and FOS, we used transcriptome analysis to look at the whole genome instead of focusing on obvious mechanisms. This allows identification of all possible biological processes active at the time of examination. Besides examination of known and expected molecular mechanisms, this technique allows the identification of new biological processes that were previously not related to the barrier function.

The aim of our microarray experiments was to identify the gene expression response of the intestinal mucosa in rats after exposure to Salmonella and FOS.

2 Main findings

1 The mucosal gene expression response to Salmonella showed that detoxification, glucose, lipid, peptide and ion transport and proteolysis were involved in the early mucosal response to Salmonella-induced stress. This was previously unreported in the literature. However, genes related to expected barrier mechanisms, such as tight junctions, mucin or toll-like receptors, were not shown to be induced. Despite clear translocation of Salmonella to extra-intestinal sites at day one post infection, the induction of immune response genes such as cytokines and chemokines was not observed at this early stage, but only later on (day 3 and day 6). This *in vivo* transcriptomic approach showed that important genes and processes, but not the expected ones, play a role in the early response of the intestinal barrier to Salmonella.

2 The early gene expression response to Salmonella in the colon was comparable in magnitude to the response observed in the ileum, indicating that Salmonella infection in rats is certainly not restricted to the small intestine, as is often presumed. The ileum and the colon show a divergent gene expression response to the same stimulus, which reflects different response mechanisms in these two intestinal segments.

3 Several genes showed an early and pronounced response to Salmonella or FOS, such as pancreatitis-associated protein, lipocalin, calprotectin and phospholipase a2. These genes are new candidates for gut health biomarkers.

4 Salmonella infection *in vivo* was found to differ from findings derived from *in vitro* systems. We observed only minor gene expression changes in the intestine following oral Salmonella infection, in contrast with *in vitro* models where major gene changes are observed. Therefore, extrapolation of results obtained in model systems to be relevant for *in vivo* gut barrier functioning should be done cautiously.

5 FOS-induced intestinal permeability was not related to the altered expression of known barrier genes such as tight junction genes, but to induced transcription of mitochondrial genes in the

colon mucosa. This finding supports a role for mitochondrial energy metabolism in barrier functioning.

6 The increased intestinal permeability induced by dietary FOS increased Salmonella translocation and coincided with enhanced induction of Salmonella-responsive genes. Increased expression of genes involved in defence or immune responses are often interpreted as increased defence and are used as surrogate markers for improved gut health. The present study shows that these surrogate markers need careful interpretation and correlation with functional effects or clinical endpoints.

3 Interpretation of findings: what determines intestinal barrier function?

3.1 Mechanisms that are part of the early gut mucosal response to Salmonella

Salmonella clearly affected detoxification in the ileum, proteolysis in the colon and transporters and antimicrobials in both segments. These processes had not previously been related to barrier functioning, but are part of general cellular maintenance. Detoxification is defined as the metabolism of xenobiotics and endogenous toxins; it involves many proteins including cytochrome P450. It is currently not understood how downregulation of detoxification affects the host response to infection. Two recent microarray studies on early barrier responses to Salmonella in pigs[30] and chickens[80] substantiate our finding that detoxification is downregulated in the early barrier response to Salmonella. Downregulation of cytochrome P450 genes expression has also been reported to occur through inflammatory mediators such as cytokines and reactive oxygen species[51,62]. However, in our study the induced expression of proinflammatory cytokines observed at day 3 p.i was mild whereas the decrease in detoxification genes at that time point was abundant. One explanation of why cells may shut down cellular systems such as detoxification in response to bacterial infection is to save energy. Another explanation is that cytochrome P450 generates epoxyeicosatrienoic acids that have anti-inflammatory effects[22]. Thus, for cells to assure progression of an inflammatory response against Salmonella, downregulation of cytochrome P450 is one mechanism to prevent formation of anti-inflammatory mediators.

It is noteworthy that the Ah receptor nuclear translocator (*Arntl*), the transcription factor regulating *Cyp1a1, Ugt1a6* and *Gsta2*[88], did show decreased expression at day 1; an additional possibility could therefore be that Salmonella actively reduces cellular detoxification systems. To clarify a role for detoxification in the host response to bacterial infection, functional studies are necessary, for example examination of protein levels, enzyme activity and histology.

In both ileal and colonic mucosa at early time points p.i., transporters and antimicrobial defence genes were altered (see Chapters 2 and 3). This differential expression of the wide range of transporters such as glucose, lipid, peptide and ion transporters may reflect increased cellular energy demand. An increase in energy demand has been reported in wounded cells. These cells increase transport of glucose into the cell to meet the metabolic demand of processes such as wound healing, cell proliferation and synthesis of extracellular matrix[75]. Since Salmonella most certainly induces cell damage of mucosal cells, this can explain the increased expression of the diverse transporters in mucosal cells. Alternatively, Salmonella might use host cellular energy for

its own purpose, ultimately leading to increased energy needs for the host cells. Additionally, the increase in transporters may be a general host response to bacterial contact, since commensal bacteria are also known to alter host transporter functioning[31]. The exact role of this process in the host barrier response requires more in-depth study.

Antimicrobial peptides are part of the host innate immune defence and are active against a broad spectrum of bacteria and other microbes[66]. Antimicrobial defence genes, such as lipocalin 2 and phospholipase A2, were among the highest induced genes after Salmonella exposure in both the ileum and colon. Antimicrobial defensin 5 was downregulated by Salmonella in both the ileum and colon. Down-regulation of defensins has also been reported previously in mucosal biopsies of patients with Shigella infection[34] and in mice orally infected with Salmonella[65]. Since Salmonella is known to manipulate host signalling[56], repression of mucosal antimicrobial peptide expression could be a sophisticated strategy of pathogenic bacteria to overcome host innate defences at the mucosal surface.

Unexpectedly, many known barrier genes, for example genes coding for mucin proteins or toll-like receptors, were not regulated by Salmonella at day 1 p.i. Immune response genes such as cytokines and chemokines were also not induced at day 1 p.i., despite clear translocation of Salmonella to extra-intestinal sites at that time point. This could indicate that Salmonella infection in rats is a targeted and controlled process. Furthermore, communication and feedback mechanisms between different mucosal cell types may help to maintain mucosal homeostasis. At later time points, days 3 and 6, immune response genes such as cytokines and chemokines were observed. However, even at these relatively late time points, induction of immune and inflammation related genes was modest.

The results of our *in vivo* transcriptomic approach supports the idea that gut barrier research should not be restricted to expected barrier processes alone. In our research, unexpected genes and processes were found to be important in the early response.

3.2 Salmonella-induced responses in different intestinal segments

In chapter 3, we showed that the colon was clearly a target for Salmonella in rats. Most Salmonella research in pigs, mice and rats focuses on the ileum and ileal Peyer's patches (PPs) as primary site of Salmonella translocation, however several reports on Salmonella infection in humans also mention involvement of the colon[49,81]. We expected involvement of the colon in Salmonella infection, since dietary FOS increases Salmonella translocation in rats[77] and FOS is known to be fermented in the colon and not in the ileum[4].

Do the ileum and colon differ in their response to Salmonella? We found that the response differed in both time and type. This variation in reactions of the ileum and colon was expected due differences in anatomy and environment. Where the ileum is composed of crypts and villi to increase the surface necessary for nutrient and water absorption[16], the surface of the colon only contains crypts, which give it a relatively smaller surfaces compared to the ileum. Furthermore, the two segments differ in environmental characteristics such as transit time and the presence of microflora. These surface and environmental differences are major determinants in the amount and the type of contact between the host and pathogens. However, general mechanisms were also identified between ileum and colon, such as changes in transporters and antimicrobials. Although both tissues showed signs of an inflammatory response at later time points, the

responsible genes were not the same. In the ileum, many chemokines were induced that were not observed in the colon. Furthermore, the response of the colon was characterized by a clear interferon γ response, which was absent in the ileum. These differences in type of immune response indicate different regulators in the two intestinal segments. This could be caused by differences in host-pathogen contact due to surface differences as mentioned above. Alternatively, the presence of microflora in the colon could cause the different colonic host response. It has been reported that host-microflora interactions keep inflammatory processes in check[12,35,40].

The magnitude of the response of colon and ileum can not be directly compared because they were studied with different array platforms. However if we compare the percentage of changed genes relative to the number of genes expressed in the specific intestinal tissues, the responses of the ileum and colon at day 1 are comparable (Figure 1). At day 3, the colon and ileum showed large difference in magnitude of gene expression responses; the colonic response was much smaller regarding the number of genes than the ileum. The PPs show an especially large response at this time point. This large response of the PPs was expected, since this tissue is part of the gut-associated lymphoid tissue and is important in activating the immune system. The delay of the colon response until day 3 could be caused by protection from the endogenous microflora, as mentioned above. This protection is very probable, since disruption of the microflora by antibiotics, chemotherapy or radiation results in increased colonization and translocation of pathogens[47]. At day 6, all three tissues responded to the same extent; the responses of all three tissues were then mainly characterized by advanced immune and inflammatory responses.
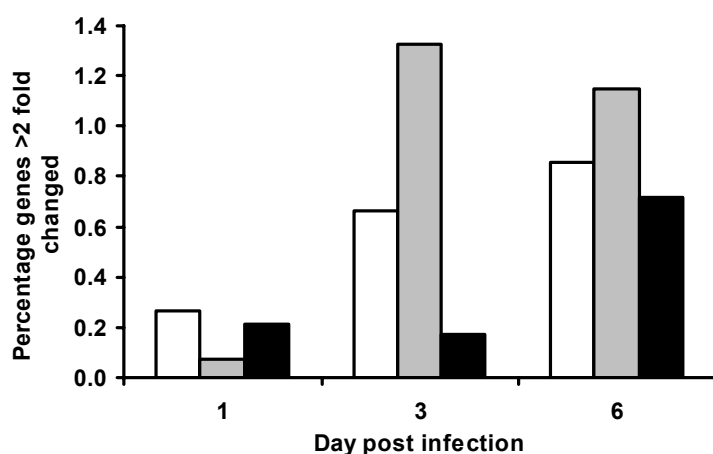
Figure 1. Percentage of genes, relative to total number of expressed genes present on the array used, that are differentially affected by Salmonella FC>2.0 in ileal mucosa (white bar), Peyer's patches (gray bar) and colonic mucosa (black bar).

3.3 Gut health biomarkers

Our approach identified several potential gut health biomarkers. A biomarker is "a characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention"[55]. Biomarkers in nutrition enable researchers to characterize and quantify the extent of a condition and the efficacy of a nutrient on the condition. Nutritional gut health research would benefit from more and better biomarkers to monitor the state of the mucosal barrier. At present, expensive and large-scale human probiotic and prebiotic studies have yielded poor results, possibly due to lack

of useful quantifiable biomarkers[61]. A biomarker can be an individual gene, protein or metabolite, or a pattern of genes, proteins or metabolites. In this research, we aimed to find individual genes. In our research, a potential new biomarker should reflect gut mucosal stress. We therefore focused on genes with a large fold change between stressed and control conditions. One candidate as a gut mucosal biomarker for Salmonella infection was Pancreatitis-associated protein (*Pap*). This gene was among the highest induced genes in both the ileum and colon. Moreover, the gene was also sensitive to dietary modulation of Salmonella infection severity, which was demonstrate by the induced expression of *Pap* in FOS fed rats (Chapter 3) and the lowered expression of *Pap* in calcium-fed rats (Chapter 4). *Pap* gene expression therefore reflected the Salmonella-induced physiological response, since FOS increases Salmonella translocation, whereas calcium inhibits Salmonella infection (Chapter 4).

Other potential biomarkers identified in Chapters 2 and 3 were lipocalin, calprotectin and phospholipase A2. Calprotectin is a promising candidate, as it is already used as a biomarker of intestinal inflammation, more specifically in monitoring inflammatory episodes and drug efficacy in IBD patients[84]. Lipocalin is also a candidate for further development, as it is highly expressed in chronically inflamed mucosa of IBD patients and in animal models of this disease[11]. Phospholipase A2 group IIA was among the highest induced genes by Salmonella and FOS, and was further induced in FOS-fed infected rats, indicating that this gene reflects the severity of gut barrier stress by different stimuli and therefore is another potential candidate biomarker.

Induced gene expression alone is not enough for a gene to function as biomarker. In addition to gene expression, the candidates should be studied at the protein level. The best candidates for nutritional research are genes for which protein products are easy to measure, preferably in biological samples obtained by non-invasive means. This enables fast and inexpensive sampling in large-scale studies with humans[58]. Easy-to-measure gene products which are secreted in stool and are resistant to enzymatic degradation by intestinal contents would be very interesting. In Chapter 4, we examined the potential gut health biomarker *Pap* at the protein level, its intestinal mucosal localization and its presence and stability in faeces. An initial step in screening for potential biomarker candidates from a large gene list could be to determine whether a gene contains a secretion signal sequence, indicating that the protein is excreted. For the purposes of gut research, excretion to the luminal site is more useful than excretion to the serosal site, as stool samples are less invasive to collect than blood samples. PAP was shown to fulfil this criterion.

Did we succeed in finding new gut health biomarkers that aid in identifying the effects of nutritional interventions? We showed that transcriptomic studies can deliver many candidates. However, studying their validity and functionality is a real effort and absolutely required in the development and application of new biomarkers. The cellular origin and faecal excretion of PAP have now been demonstrated. But additional studies are required, for example to find out whether the protein behaves similarly in humans. A final bottleneck for the application of new biomarkers is that golden standard biomarkers are not easily replaced. Acceptance of new candidates requires many published studies showing good correlation between old and new biomarkers. We offered new ideas, but further validation studies are required to bring these markers to the point that they will be accepted as a read-out of the physiological state of the intestinal mucosa.

3.4 *In vivo* gut barrier versus *in vitro* models

The early *in vivo* modulation of gene expression by Salmonella was small in both the ileum and colon. This contrasts with the results reported in infection studies of *in vitro* cell models, where pathogens cause much larger gene expression changes[19]. Apparently, the gut mucosal cells in the intact barrier display much more subtle responses. One explanation for this subtle response in the intact tissue is that different cell types interact together via various feedback mechanisms that keep the reaction against invading pathogens in check. This differs from *in vitro* cell cultures consisting of only one cell type. Furthermore, in the intact animal, the protective mucus layer and microflora provide an effective barrier to pathogens. Therefore bacterial-host contact is probably less than in cell cultures. The mild gene expression changes *in vivo* are in accordance with physiological observations. No inflammatory changes are observed *in vivo* on the first day after oral infection with Salmonella, whereas *in vitro* models show massive cell death after 24 hours of Salmonella exposure and ileal loop models show epithelial detachment after 8 hours of exposure to Salmonella[67].

Besides physiological changes, the small gene expression changes in our study could be partly explained by the technical approach used. The collected mucosal scrapings contain a heterogeneous population of cells, whereas *in vitro* cell cultures contain only one cell type, and all cells are exposed to the pathogen at the same time point, often at a high dosage. As a result, genes expressed in only a subset of cells in the mucosal scrapings might be missed due to dilution of the signal. This mixture and possible dilution of cells can be overcome by using laser capture techniques. This technique has proven to be successful in isolating specific cell types from intestinal tissue[72]. Although the RNA yield per tissue is much lower using this technique, with the rapid progress in microarrays using amplification, arrays can already be performed from 50 nanograms of total RNA, thus enabling this cell-specific approach.

When compared to data from *in vitro* studies, the results of our *in vivo* approach show that the intact gut barrier system must be examined in order to study barrier responses to bacterial or nutritional stress. *In vitro* models are useful for in-depth studies of one particular cell type or process. Similarly, intestinal loop models can be useful for screening several components or products, because this method allows the use of several loops per animal, thereby decreasing animal numbers and variation between samples. In both model systems, however, extrapolation to the *in vivo* relevance for gut barrier functioning should be done cautiously.

3.5 Dietary FOS-induced gut mucosal response

Despite FOS-induced intestinal permeability, we did not observe changes in tight junction genes as expected. However, FOS induced a large number of mitochondrial genes in the colon mucosa. Increased expression of mitochondrial genes indicates disturbed energy homeostasis, or more specifically ATP homeostasis, caused by increased energy demand or decreased mitochondrial energy supply[60,86]. We showed for the first time that the increased barrier permeability induced by FOS supplementation coincides with alterations in genes related to energy metabolism. Previously, a relation between increased intestinal permeability and intestinal energy metabolism was observed in studies with chronic supplementation of non-steroidal anti-inflammatory drugs or administration of mitochondrial uncouplers to animals[64]. Together with our results, this suggests that energy metabolism is important in barrier functioning and is likely involved in the

FOS-induced barrier changes. The exact role of energy metabolism in FOS-induced mucosal effects requires additional experimental proof. The next important step would be to measure the energy status of intestinal cells by measuring energy metabolism parameters such as the ratios of cellular ATP to ADP and AMP[64]. Additional proof, that decreased ATP levels cause the FOS-induced intestinal permeability, can be derived from studies on direct ATP supplementation, similar to that of Bours et al. In that study, local intestinal co-administration of ATP with non-steroidal anti-inflammatory drugs attenuated the induced intestinal permeability[8]. It would be interesting to examine whether ATP supplementation to the colon of FOS-fed rats would lead to decreased permeability. Further experimental proof of the suggested FOS effects on mitochondria could also come from electron microscopy studies, since in uncoupling studies in fibroblasts, Rohas et al showed that upregulation of genes encoding for mitochondrial proteins coincided with increased mitochondrial density[64]. Mitochondrial swelling, ATP depletion, and permeability changes are also observed in other situations of weakened gut barrier, for example in mucosal biopsies of Crohn's disease patients and in intestinal tissue after ischemia-reperfusion[43,63,71]. Additional studies using direct administration of uncouplers in the intestine could clarify whether altered energy metabolism is also causing the FOS-induced increased mucus secretion and decreased resistance to Salmonella.

FOS is currently being studied in the obesity research field because it induces release of gastrointestinal peptides such as glucagon-like protein, which is thought to promote satiety[10]. The use of mitochondrial uncouplers is also being studied as a possible therapeutic approach to treat obesity, since mild but chronic treatment with mitochondrial uncouples should cause increased energy expenditure[64]. Our results indicate that the potential negative effects of these two approaches on the gut barrier should be given attention in both these research fields.

3.6 FOS effect on Salmonella-induced colonic gene expression

In FOS-fed rats, three times more genes were significantly altered by Salmonella compared to cellulose-fed control rats. This colonic gene expression response reflects the increase in physiological parameters that are known to be modulated by FOS during Salmonella infection, such as Salmonella translocation, mucin secretion and intestinal permeability in rats[77]. Although the magnitude of the gene expression response was larger in FOS-fed rats infected with Salmonella, the genes reflected the same biological processes in both FOS-fed and cellulose-fed rats. This could indicate that the intestine uses generic mechanisms to respond to a variety of harmful stimuli. Oxidative stress and proteolysis were among the biological processes that were induced in the colonic mucosa by both Salmonella (Chapter 3) and FOS (Chapter 6). Indeed, increased oxidative stress can cause protein oxidation and subsequent misfolding, leading to increased protein degradation[37]. An alternative explanation for the increased proteolysis is the need for increased antigen presentation in the stressed mucosa, which is known to require protein breakdown by immunoproteasomes[83]. Additional transcriptomic studies of the intestine under healthy conditions and a variety of stressed conditions are necessary to distinguish generic mechanisms from stress stimuli-dependent mechanisms.

The potential beneficial effects of FOS and other prebiotics are based on effects of these nutrients on promotion of surrogate markers, such as increased mucin production, increased cytokine production and increased faecal or ileal IgA[32,52,69,74,85]. Changes in these markers are often

presumed to reflect increased barrier protection or resistance to pathogenic bacteria, but actual measurement of these functional effects is lacking in most studies. An increased immune response activation might however not indicate beneficial increased surveillance, but could actually indicate induced tissue damage. In our study, genes involved in antimicrobial defence, immune response and inflammation were all induced by Salmonella infection and further enhanced by dietary FOS. This coincided with increased translocation of Salmonella by dietary FOS. Thus, these surrogate markers should be interpreted with caution and always interpreted in conjunction with functional effects or clinical endpoints.

4 Transcriptomics in gut barrier research: technical aspects and data analysis

Transcriptomics is thought to be most successfully used in a hypothesis-driven approach, but we have shown that this technique is also successful in hypothesis generation. Specifically, we used information provided by differential gene expression caused by Salmonella and FOS to generate new hypotheses about the gut mucosal barrier. The generated hypothesis resulted in leads for further gut barrier research associated with energy metabolism, detoxification and transporters. However, such hypothesis-generating research does require a clear study design to prevent a subjective search for genes which could be selected by chance and are actually noise variables. Distinguishing between true differences and noise in microarray datasets is a challenging task, especially in gene expression datasets derived from nutritional studies, which are characterized by subtle changes and large variability between the individuals studied[1,14,57]. Indeed, both the dietary FOS dataset and the Salmonella dataset were characterized by subtle gene expression changes.

The transcriptomic studies conducted in this thesis research delivered several technical aspects regarding experimental set up and data analysis. These aspects are important to extract reliable and valuable information from microarray datasets with subtle gene expression and to convert gene expression data into biological insights.

4.1 Reproducibility

In general, the gene expression data were highly reproducible between experiments and array platforms used in the studies. Three array platforms, cDNA, oligo and commercial, were used in the research, two of which, oligo and commercial, are described in this thesis. Reported differences between platforms have raised doubts about the quality and reproducibility of microarrays[44,59]. However, with solid experimental design and strict quality control, all platforms can offer qualitatively good data. This was demonstrated by a large microarray quality control project that showed consistent results between different platforms and laboratories[70]. This large study proved the reliability and reproducibility of microarrays technology for assessing gene expression changes. In agreement with this study, we observed that different platforms, independent hybridizations or independent biological samples showed consistent data. For example, the Salmonella-induced increase of the gene *Pap* was identified on all three platforms. Furthermore, the results from all three platforms were consistently confirmed by an independent method: quantitative real-time RT-PCR (Q-PCR).

4.2 Experimental set up

Q-PCR confirmation of the pooled array data in Chapters 2 and 3 showed that the inter-individual variation between the rats was high. The selection method based on fold change, does not take this variability in the data into account. In contrast, datasets derived from individual hybridization provide insight into this variability and therefore enable inclusion of subtle but statistically significant data. Using this insight into the inter-individual variation in gene expression in the rats, we choose a different approach for the dietary FOS study (Chapters 5 and 6). In this study we enlarged the group size from 8 to 12 rats per group to obtain more power for detection of significant treatment effects, and we performed the micorrarray experiment on individual rats instead of using pooled samples. This allowed the use of statistical tests to analyze the data instead of fold change alone.

It is recognized that in micorarray datasets true differences are difficult to distinguish from noise; this is not only caused by the large variability among individuals studied or the small sample sizes, but especially by the large number of genes tested[50,90]. This is because testing these large numbers of genes, up to 40.000, increases the chance of selecting false positive genes due to the multiple testing problem. For example, by testing a dataset of 10.000 genes, a probability of 5% results in selection of about 500 genes just by chance. Therefore, an important aspect in data analysis is the choice of an appropriate statistical test.

4.3 Statistical data analysis

The growing popularity of transcriptomics coincides with a growing list of statistical tools to analyze transcriptomic data[33], for example multiple testing procedures (FDR, FWER)[6,7], significance analysis of microarray data (SAM)[79], analysis of variance (ANOVA), empirical bayes t-statistic and many others. There are several drawbacks to most of these statistical methods. First, most of the methods were developed for classification, which intends to select the smallest set of most differential genes. Second, the methods are tested on gene sets consisting of large gene expression differences such as cancer studies or pharmaceutical interventions. Third, most methods consider genes as independent variables, which means that each gene is tested independently. These tests can successfully identify the main effects in a dataset. However, genes are not completely independent. Small genetic effects that make an important distinction between different study groups only in interaction with each other, will therefore not be detected with these methods. These genes should be included in the selection, especially in nutritional studies, which are characterized by subtle gene expression effects. These interactions between genes are important in biology because gene-gene interactions occur during biological responses due to co-regulation of genes, for example by one transcription factor[29]. It is also recognized that genes with similar functions often share similar overall expressions due to common activators or inhibitors[20,46,87]. Moreover, genes that encode different subunits of larger complexes are coordinately regulated[50]. Therefore, statistical tests should not only test each gene independently, but preferably include interaction between genes. In Chapter 5 we applied the Random Forest method, which integrates interaction into the statistical analysis. Random Forests not only takes interaction into account, but also ranks genes. We examined whether Random Forests could provide an objective threshold to distinguish real effects from noise. Choosing a threshold is

difficult but important, since the threshold choice determines the selection of relevant genes and therefore the translation into biological information. In microarray datasets obtained from pooled samples, fold change is often used as threshold, with the precise cut-off arbitrarily chosen by the researcher. The same is true for datasets obtained from individual hybridizations, where thresholds for statistical tests are also arbitrarily chosen to some extent. With Random Forests, we were able to define an objective threshold for gene selection. Biological interpretation of these selected genes is the next step.

4.4 Functional data analysis

Microarrays have grown in size from several hundred genes on cDNA arrays to about 40.000 genes on the commercial whole genome arrays. As a result, manual searching for biological interpretation of selected gene lists became difficult and time consuming. Pathway analysis is an increasingly popular and useful way to interpret large microarray derived gene lists. Pathway analysis programs group individual gene lists into functional biological categories, such as biological processes, signal transduction pathways or cellular location[21]. The number of pathway programs has expanded rapidly over the last few of years, including open source programs such as David[15], ErmineJ[45] and GSEA[73] as well as commercial programs such as Metacore[21] and Ingenuity. All programs use the same basic principles; they require the input of a list of gene identifiers such as accession number or gene symbol, and they link this list to gene annotation databases such as Gene Ontology[25], Kyoto Encyclopedia of Genes and Genomes (KEGG)[38] or Fatigo[3], returning groups of genes that belong to the same functional class (reviewed in Dopazo et al[17]). The programs calculate if certain functional classes are statistically overrepresented in the selected gene list compared to the rest of the genome[28]. This information provides objective insight into biological processes that are responding to the treatment.

One advantage of pathway analysis is that it allows reliable identification of subtle changes in gene expression. Especially in studies with overall subtle gene expression changes, analysis on the level of biological processes, instead of individual gene lists, reduces the chance of introducing false positive genes[13]. This is because genes are not tested individually, but in the context of a functional group. Genes with subtle gene expression differences within one significantly regulated biological process can be selected with greater certainty than similarly subtle changed genes not belonging to an altered biological process. This is essential in nutritional studies, and as shown in Chapters 2 and 3, also for early events of pathogen-host responses. In those two chapters we performed pathway analysis with all genes changed more than twofold, but added genes with a fold change between 1.5 and twofold when these genes were part of a selected biological process, since this decreased the probability that these genes were selected by chance.

Another advantage of analysis at the biological process level is that experiments are more comparable at the process level than at the level of individual gene lists. This is shown in Chapter 3, where two independent Salmonella infection experiments of the colon mucosa showed highly similar results at the biological process level, but at the individual gene level these two independent studies gave only moderate overlap. The improved comparison also accounts for different studies. For example, in Chapter 3 and Chapter 6, proteolysis was found to be induced

in the colon by Salmonella and by FOS respectively. In addition, Salmonella-induced decrease of detoxification was found in our rat studies and recently in chickens[80] and pigs[30], although these studies used different array platforms and individual genes did not fully overlap. This lack of overlap at individual gene level and large overlap at process level is inherent to mRNA studies; genes are switched on and off quickly and the half-life of mRNA is short. However, genes within a similar process are likely to be similarly regulated and therefore more overlap at the process level will occur than at individual gene expression level[2]. This clearly highlights the importance of interpreting experiments at the level of biological processes rather than using gene-to-gene comparison.

However, a disadvantage of complete dependency on pathway programs is that these programs rely on the ability to assign functions to each regulated gene[41]. Annotation of genes and uniform naming is far from optimal. The availability of genome information for humans, mice, rats and other species has accelerated the identification of genes involved in biological processes, but it is important to realize that for many genes, no functional information is yet available. About half of the probes on the commercial whole genome arrays represent transcripts of unknown function and are therefore excluded from further pathway analysis, despite possible significant regulation (Chapter 5). Pathway analysis is therefore not fully functional. Another problem is microarray probes ID. There is no single database available that can cross reference between various gene identifiers such as accession number, unigene number, agilent ID, gene symbol and protein name. This is problematic and might cause loss of genes during the analysis, since input ID does not match the ID of the database used by the pathway program[41]. Functional annotation and uniform naming is being given increasing attention and will most likely improve rapidly. Additionally, pathway programs are increasingly incorporating refinements to construct regulatory networks and processes not only based on biological function, but also on other types of annotations such as promoter elements, chromosome position or sequence information about secretion signals[23,42]. This will increase biological insight into microarray data and will show that these long gene lists are very suitable to study physiology.

4.5 Connecting array data to physiology

Array data should preferably be linked to strong physiological data, such as intestinal permeability or Salmonella translocation. Otherwise it is difficult to determine whether changes in gene expression are positive or negative for the organism. In Chapter 6, the data on FOS increased permeability were essential to link the effects on mitochondrial genes to negative effects on the mucosal barrier. And in Chapter 3, the known increase in Salmonella translocation in FOS-fed rats was required to interpret the enhanced gene expression data in FOS-fed vs. cellulose-fed infected rats as a negative effect. In this study we have shown that transcriptomics is extremely suitable as a hypothesis generating technique, provided that physiological data are available. These two aspects, transcriptomics and physiology, should be combined much more strongly to increase insight into the molecular mechanisms that link the treatment, for instance FOS, with the physiological output, for instance intestinal permeability.

Microarray data as new complementary data for physiology is not always accepted. This is mainly due to early failures in generating useful data[48]. However, with the improved quality and reliability of microarrays, the growing insight into data analysis and the availability of comparative datasets, acceptance should improve.

5 Recommendations

5.1 Follow-up studies

The transcriptomic approach we used in our research provided new hypotheses about the response of the intestinal barrier to stress. Processes likely to be involved in barrier functioning are detoxification, transport, energy metabolism and proteolysis. However, mRNA expression does not necessarily reflect protein expression and activity; mRNA stability, post-translational modification, protein localization and protein degradation determine whether a gene is translated to a protein and whether the protein is active and metabolites are produced. Not all genes are ultimately transcribed into proteins, and not all transcribed proteins are active. Secondary modulation, such as phosphorylation, acetylation or cleavage, is often necessary for full activity[9]. Hypotheses on biological processes or possible biomarkers observed in microarray studies should therefore be followed up with more detailed or functional studies to really define a role for these biomarkers and processes in gut barrier function. Confirmatory studies on protein level (as in Chapter 4) are the most logical follow-up studies for potential biomarkers. To resolve the role of one gene in the *in vivo* situation knock outs, RNAi mediated gene silencing or genetically modified animal models can be used. Recently, a PAP knock-out mouse was developed, and it would be very interesting to study the resistance of these mice to Salmonella[27]. More in-depth studies on the role of energy metabolism on gut barrier function are also required, such as functional studies to define the role of ATP in barrier functioning or protein studies which can reveal possible effects of FOS on regulation of tight junction proteins. Although FOS did not affect tight junction genes, immuno-histochemistry could show effects on cellular re-distribution of tight junction proteins affecting paracellular permeability. Another field that is gaining more attention is the combination of transcriptomic with other broad-scale techniques such as proteomics and metabolomics[24,78]. This combination will provide information about the more downstream effects of a treatment. These and additional functional studies will help to provide further insight into gut barrier physiology. Another approach to gain insight into the newly found barrier mechanisms is comparison with other gene expression datasets derived from *in vivo* barrier studies. As mentioned earlier, detoxification was recently related to barrier function in other species than rats. This shows that comparison of the microarray results in our study with the results of other related studies lead to confirmation of one observed barrier process, detoxification, and might reveal more overlapping mechanisms. The strength of this large comparison approach is shown by Jenner et al.[36], who compared 160 *in vitro* transcriptomic datasets of the host response to various pathogens. This revealed a general host response to pathogens. Similar comparisons of *in vivo* data might identify new host response processes as possible targets for dietary or pharmaceutical modulation. To enable comparisons of different studies, datasets should become publicly available. An important contribution in this regard is

that most journals now require all microarray datasets to be deposited in public databases such as the Gene expression omnibus[54] or Arrayexpress[5].

5.2 Application of findings for humane gut research

It is not possible to directly extrapolate the results obtained in rats to the human situation. The rat model is a good model for Salmonella infection and dietary studies. Nevertheless, rats are not humans. The rat intestine differs from the human intestine; the cecum of rodents is relatively much larger than in humans. The relative length of the cecum compared to the total length of the large intestine is 26% in rats and only 5% in humans[16]. Although the cecum is involved in FOS fermentation in rats, we decided to focus on FOS effects in the rat colon, as the large intestine is more relevant to fermentation of prebiotics in humans.

Another difficulty with extrapolation is that gene expression patterns in one strain, or genotype, may be unique to that genotype[39,89]. To minimize genotype-dependent results, we chose to use an outbred rat strain[82]. The use of outbred rats requires larger study groups than inbred strains to obtain equal power for detection of significant treatment effects. In Chapter 6, subtle differences were identified as highly significant using 12 animals per group. Human studies cannot be controlled as rigidly as animal studies, and most likely should include many more subjects to reach enough power. However, several aspects of our transcriptomic approach are relevant for human gut research. For example, statistical methods that test gene-gene interaction could also be more powerful in human genomics studies, where individual variation might cover up results when each gene is tested individually. Nevertheless animal studies are relevant to the human situation, since these studies can deliver knowledge on treatments that are impossible to obtain from human studies, for example insight into dietary modulation of a Salmonella infection. Furthermore, animal studies allow much easier sampling of biological material such as intestinal tissue than human studies. These biological materials are required to identify new biomarkers, which can subsequently be used in human studies.

6 Conclusions

We determined that detoxification, transport, energy metabolism and proteolysis are part of the early gene expression responses of the rat gut mucosa upon exposure to Salmonella and FOS. Expected barrier related genes were not affected by Salmonella or FOS, and immune response genes showed a delayed and mild response. Barrier research should therefore not only focus on a small panel of "known" barrier genes. Furthermore, our data clearly showed that the barrier in an intact organism shows much milder responses following Salmonella challenge than *in vitro* models. Therefore, extrapolating data from model systems to the intact animal should be done cautiously.

Analyzing transcriptomic data with subtle gene expression changes is a powerful approach, provided that a solid strategy is followed to prevent discarding valuable data or introducing noise. Before newly found processes and biomarkers that are identified in transcriptome analysis can be used in practice, functional follow-up studies are required. Nevertheless, our transcriptome analyses have significantly increased the understanding of the actual *in vivo* barrier processes and have provided starting points for functional studies in gut barrier research.

References

1. Afman, L. & Muller, M. Nutrigenomics: from molecular nutrition to prevention of disease. *J Am Diet Assoc* 106, 569-576 (2006).
2. Al-Shahrour, F.*, et al.* From genes to functional classes in the study of biological systems. *BMC bioinformatics* 8, 114 (2007).
3. Al-Shahrour, F., Diaz-Uriarte, R. & Dopazo, J. FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes. *Bioinformatics* 20, 578-580 (2004).
4. Andersson, H.B., Ellegard, L.H. & Bosaeus, I.G. Nondigestibility characteristics of inulin and oligofructose in humans. *J Nutr* 129, 1428S-1430S (1999).
5. ArrayExpress, E.-E. http://www.ebi.ac.uk/microarray-as/aer/?#ae-main[0].
6. Benjamini, Y. & Yekutieli, D. The control of the false discovery rate in multiple testing under dependency. *Ann Statist* 29, 1165-1188 (2001).
7. Benjamini, Y.H., Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc.* 57, 289-300 (1995).
8. Bours, M.J., Troost, F.J., Brummer, R.J., Bast, A. & Dagnelie, P.C. Local effect of adenosine 5'-triphosphate on indomethacin-induced permeability changes in the human small intestine. *Eur J Gastroenterol Hepatol* 19, 245-250 (2007).
9. Brooks, C.L. & Gu, W. Ubiquitination, phosphorylation and acetylation: the molecular basis for p53 regulation. *Current opinion in cell biology* 15, 164-171 (2003).
10. Cani, P.D., Joly, E., Horsmans, Y. & Delzenne, N.M. Oligofructose promotes satiety in healthy human: a pilot study. *European journal of clinical nutrition* 60, 567-572 (2006).
11. Carlson, M.*, et al.* Human neutrophil lipocalin is a unique marker of neutrophil inflammation in ulcerative colitis and proctitis. *Gut* 50, 501-506 (2002).
12. Collier-Hyams, L.S., Sloane, V., Batten, B.C. & Neish, A.S. Cutting edge: bacterial modulation of epithelial signaling via changes in neddylation of cullin-1. *J Immunol* 175, 4194-4198 (2005).
13. Curtis, R.K., Oresic, M. & Vidal-Puig, A. Pathways to the analysis of microarray data. *Trends in biotechnology* 23, 429-435 (2005).
14. de Boer, V.C.*, et al.* Chronic quercetin exposure affects fatty acid catabolism in rat lung. *Cell Mol Life Sci* 63, 2847-2858 (2006).
15. Dennis, G., Jr.*, et al.* DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome biology* 4, P3 (2003).
16. DeSesso, J.M. & Jacobson, C.F. Anatomical and physiological parameters affecting gastrointestinal absorption in humans and rats. *Food Chem Toxicol* 39, 209-228 (2001).
17. Dopazo, J. Functional interpretation of microarray experiments. *Omics* 10, 398-410 (2006).
18. Duggan, C., Gannon, J. & Walker, W.A. Protective nutrients and functional foods for the gastrointestinal tract. *The American journal of clinical nutrition* 75, 789-808 (2002).
19. Eckmann, L., Smith, J.R., Housley, M.P., Dwinell, M.B. & Kagnoff, M.F. Analysis by high density cDNA arrays of altered gene expression in human intestinal epithelial cells in response to infection with the invasive enteric bacteria Salmonella. *J Biol Chem* 275, 14084-14094 (2000).
20. Eisen, M.B., Spellman, P.T., Brown, P.O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 95, 14863-14868 (1998).
21. Ekins, S., Nikolsky, Y., Bugrim, A., Kirillov, E. & Nikolskaya, T. Pathway mapping tools for analysis of high content data. *Methods Mol Biol* 356, 319-350 (2007).
22. Fleming, I. DiscrEET regulators of homeostasis: epoxyeicosatrienoic acids, cytochrome P450 epoxygenases and vascular inflammation. *Trends in pharmacological sciences* 28, 448-452 (2007).
23. Gao, F., Foat, B.C. & Bussemaker, H.J. Defining transcriptional networks through integrative modeling of mRNA expression and transcription factor binding data. *BMC bioinformatics* 5, 31 (2004).
24. Ge, H., Walhout, A.J. & Vidal, M. Integrating 'omic' information: a bridge between genomics and systems biology. *Trends Genet* 19, 551-560 (2003).
25. Gene Ontology Consortium. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res* 32, D258-261 (2004).
26. Gewirtz, A.T., Liu, Y., Sitaraman, S.V. & Madara, J.L. Intestinal epithelial pathobiology: past, present and future. *Best practice & research* 16, 851-867 (2002).
27. Gironella, M.*, et al.* Experimental acute pancreatitis in PAP/HIP knock-out mice. *Gut* 56, 1091-1097 (2007).
28. Goeman, J.J. & Buhlmann, P. Analyzing gene expression data in terms of gene sets: methodological issues. *Bioinformatics* 23, 980-987 (2007).

29. Hallikas, O., *et al.* Genome-wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. *Cell* 124, 47-59 (2006).

30. Handley, S.A. & Miller, V.L. General and specific host responses to bacterial infection in Peyer's patches: a role for stromelysin-1 (matrix metalloproteinase-3) during Salmonella enterica infection. *Molecular microbiology* 64, 94-110 (2007).

31. Hooper, L.V., *et al.* Molecular analysis of commensal host-microbial relationships in the intestine. *Science (New York, N.Y* 291, 881-884 (2001).

32. Hosono, A., *et al.* Dietary fructooligosaccharides induce immunoregulation of intestinal IgA secretion by murine Peyer's patch cells. *Bioscience, biotechnology, and biochemistry* 67, 758-764 (2003).

33. Hurley, B.P. & McCormick, B.A. Translating tissue culture results into animal models: the case of Salmonella typhimurium. *Trends in microbiology* 11, 562-569 (2003).

34. Islam, D., *et al.* Downregulation of bactericidal peptides in enteric infections: a novel immune escape mechanism with bacterial DNA as a potential regulator. *Nature medicine* 7, 180-185 (2001).

35. Ismail, A.S. & Hooper, L.V. Epithelial cells and their neighbors. IV. Bacterial contributions to intestinal epithelial barrier integrity. *Am J Physiol Gastrointest Liver Physiol* 289, G779-784 (2005).

36. Jenner, R.G. & Young, R.A. Insights into host responses against pathogens from transcriptional profiling. *Nature reviews* 3, 281-294 (2005).

37. Jung, T., Bader, N. & Grune, T. Oxidized proteins: intracellular distribution and recognition by the proteasome. *Arch Biochem Biophys* 462, 231-237 (2007).

38. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 28, 27-30 (2000).

39. Kaput, J. & Rodriguez, R.L. Nutritional genomics: the next frontier in the postgenomic era. *Physiological genomics* 16, 166-177 (2004).

40. Kelly, D., *et al.* Commensal anaerobic gut bacteria attenuate inflammation by regulating nuclear-cytoplasmic shuttling of PPAR-gamma and RelA. *Nature immunology* 5, 104-112 (2004).

41. Khatri, P. & Draghici, S. Ontological analysis of gene expression data: current tools, limitations, and open problems. *Bioinformatics* 21, 3587-3595 (2005).

42. Kim, S.Y. & Kim, Y. Genome-wide prediction of transcriptional regulatory elements of human promoters using gene expression and promoter analysis data. *BMC bioinformatics* 7, 330 (2006).

43. Kreienberg, P.B., Darling, R.C., 3rd, Shah, D.M., Vincent, P.A. & Blumenstock, F.A. ATP-MgCL2 reduces intestinal permeability during mesenteric ischemia. *The Journal of surgical research* 66, 69-74 (1996).

44. Larkin, J.E., Frank, B.C., Gavras, H., Sultana, R. & Quackenbush, J. Independence and reproducibility across microarray platforms. *Nature methods* 2, 337-344 (2005).

45. Lee, H.K., Braynen, W., Keshav, K. & Pavlidis, P. ErmineJ: tool for functional analysis of gene expression data sets. *BMC bioinformatics* 6, 269 (2005).

46. Lee, H.K., Hsu, A.K., Sajdak, J., Qin, J. & Pavlidis, P. Coexpression analysis of human genes across many microarray data sets. *Genome research* 14, 1085-1094 (2004).

47. Levy, J. The effects of antibiotic use on gastrointestinal function. *The American journal of gastroenterology* 95, S8-10 (2000).

48. Liang, M., Cowley, A. & Greene, A. High throughput gene expression profiling: a molecular approach to integrative physiology. *J Physiol* 554, 22-30 (2004).

49. Mandal, B.K. & Mani, V. Colonic involvement in salmonellosis. *Lancet* 1, 887-888 (1976).

50. Mootha, V.K., *et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 34, 267-273 (2003).

51. Morgan, E.T., Li-Masters, T. & Cheng, P.Y. Mechanisms of cytochrome P450 regulation by inflammatory mediators. *Toxicology* 181-182, 207-210 (2002).

52. Nakamura, Y., *et al.* Dietary fructooligosaccharides up-regulate immunoglobulin A response and polymeric immunoglobulin receptor expression in intestines of infant mice. *Clinical and experimental immunology* 137, 52-58 (2004).

53. Naughton, P.J., Grant, G., Spencer, R.J., Bardocz, S. & Pusztai, A. A rat model of infection by Salmonella typhimurium or Salm. enteritidis. *The Journal of applied bacteriology* 81, 651-656 (1996).

54. NCBI GEO website. http://www.ncbi.nlm.nih.gov/geo/. (,).

55. NIH Definition Working Group. Biomarkers and surrogate endpoints in clinical research: defintions and conceptual model. in *Biomarkers and Surrogate Endpoints: Clinical Research and Applications* (ed. Downing, G.) 1-9 (Elsevier, Amsterdam, 2000).

56. Patel, J.C. & Galan, J.E. Differential activation and function of Rho GTPases during Salmonella-host cell interactions. *The Journal of cell biology* 175, 453-463 (2006).

57.   Patsouris, D., Reddy, J.K., Muller, M. & Kersten, S. Peroxisome proliferator-activated receptor alpha mediates the effects of high-fat diet on hepatic gene expression. *Endocrinology* 147, 1508-1516 (2006).

58.   Poullis, A., Foster, R., Northfield, T.C. & Mendall, M.A. Review article: faecal markers in the assessment of activity in inflammatory bowel disease. *Aliment Pharmacol Ther* 16, 675-681 (2002).

59.   Quackenbush, J. Computational analysis of microarray data. *Nat Rev Genet* 2, 418-427 (2001).

60.   Rachamim, N.*, et al.* Dexamethasone enhances expression of mitochondrial oxidative phosphorylation genes in rat distal colon. *Am J Physiol* 269, C1305-1310 (1995).

61.   Rafter, J.*, et al.* Dietary synbiotics reduce cancer risk factors in polypectomized and colon cancer patients. *The American journal of clinical nutrition* 85, 488-496 (2007).

62.   Renton, K.W. Alteration of drug biotransformation and elimination during infection and inflammation. *Pharmacol Ther* 92, 147-163 (2001).

63.   Restivo, N.L., Srivastava, M.D., Schafer, I.A. & Hoppel, C.L. Mitochondrial dysfunction in a patient with crohn disease: possible role in pathogenesis. *Journal of pediatric gastroenterology and nutrition* 38, 534-538 (2004).

64.   Rohas, L.M.*, et al.* A fundamental system of cellular energy homeostasis regulated by PGC-1alpha. *Proc Natl Acad Sci U S A* 104, 7933-7938 (2007).

65.   Salzman, N.H.*, et al.* Enteric salmonella infection inhibits Paneth cell antimicrobial peptide expression. *Infection and immunity* 71, 1109-1115 (2003).

66.   Salzman, N.H., Underwood, M.A. & Bevins, C.L. Paneth cells, defensins, and the commensal microbiota: a hypothesis on intimate interplay at the intestinal mucosa. *Seminars in immunology* 19, 70-83 (2007).

67.   Santos, R.L., Zhang, S., Tsolis, R.M., Baumler, A.J. & Adams, L.G. Morphologic and molecular characterization of Salmonella typhimurium infection in neonatal calves. *Vet Pathol* 39, 200-215 (2002).

68.   Santos, R.L.*, et al.* Animal models of Salmonella infections: enteritis versus typhoid fever. *Microbes and infection / Institut Pasteur* 3, 1335-1344 (2001).

69.   Schmidt-Wittig, U., Enss, M.L., Coenen, M., Gartner, K. & Hedrich, H.J. Response of rat colonic mucosa to a high fiber diet. *Annals of nutrition & metabolism* 40, 343-350 (1996).

70.   Shi, T., Seligson, D., Belldegrun, A.S., Palotie, A. & Horvath, S. Tumor classification by tissue microarray profiling: random forest clustering applied to renal cell carcinoma. *Mod Pathol* 18, 547-557 (2005).

71.   Soderholm, J.D.*, et al.* Augmented increase in tight junction permeability by luminal stimuli in the non-inflamed ileum of Crohn's disease. *Gut* 50, 307-313 (2002).

72.   Stappenbeck, T.S., Hooper, L.V., Manchester, J.K., Wong, M.H. & Gordon, J.I. Laser capture microdissection of mouse intestine: characterizing mRNA and protein expression, and profiling intermediary metabolism in specified cell populations. *Methods in enzymology* 356, 167-196 (2002).

73.   Subramanian, A.*, et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102, 15545-15550 (2005).

74.   Swanson, K.S.*, et al.* Supplemental fructooligosaccharides and mannanoligosaccharides influence immune function, ileal and total tract nutrient digestibilities, microbial populations and concentrations of protein catabolites in the large bowel of dogs. *J Nutr* 132, 980-989 (2002).

75.   Takahashi, H., Kaminski, A.E. & Zieske, J.D. Glucose transporter 1 expression is enhanced during corneal epithelial wound repair. *Experimental eye research* 63, 649-659 (1996).

76.   Ten Bruggencate, S.J., Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L. & Van der Meer, R. Dietary fructo-oligosaccharides dose-dependently increase translocation of salmonella in rats. *J Nutr* 133, 2313-2318 (2003).

77.   Ten Bruggencate, S.J., Bovee-Oudenhoven, I.M., Lettink-Wissink, M.L. & Van der Meer, R. Dietary fructooligosaccharides increase intestinal permeability in rats. *J Nutr* 135, 837-842 (2005).

78.   Thomas, C.E. & Ganji, G. Integration of genomic and metabonomic data in systems biology--are we 'there' yet? *Current opinion in drug discovery & development* 9, 92-100 (2006).

79.   Tusher, V.G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 98, 5116-5121 (2001).

80.   van Hemert, S., Hoekman, A.J., Smits, M.A. & Rebel, J.M. Gene expression responses to a Salmonella infection in the chicken intestine differ between lines. *Veterinary immunology and immunopathology* 114, 247-258 (2006).

81.   Vender, R.J. & Marignani, P. Salmonella colitis presenting as a segmental colitis resembling Crohn's disease. *Dig Dis Sci* 28, 848-851 (1983).

82.   Wade, C.M.*, et al.* The mosaic structure of variation in the laboratory mouse genome. *Nature* 420, 574-578 (2002).

83.   Wang, J. & Maldonado, M.A. The ubiquitin-proteasome system and its role in inflammatory and autoimmune diseases. *Cell Mol Immunol* 3, 255-261 (2006).

84. Wassell, J., Dolwani, S., Metzner, M., Losty, H. & Hawthorne, A. Faecal calprotectin: a new marker for Crohn's disease? *Ann Clin Biochem* 41, 230-232 (2004).

85. Watzl, B., Girrbach, S. & Roller, M. Inulin, oligofructose and immunomodulation. *The British journal of nutrition* 93 Suppl 1, S49-55 (2005).

86. Wiesner, R.J.*, et al.* Stimulation of mitochondrial gene expression and proliferation of mitochondria following impairment of cellular energy transfer by inhibition of the phosphocreatine circuit in rat hearts. *J Bioenerg Biomembr* 31, 559-567 (1999).

87. Wolfe, C.J., Kohane, I.S. & Butte, A.J. Systematic survey reveals general applicability of "guilt-by-association" within gene coexpression networks. *BMC bioinformatics* 6, 227 (2005).

88. Xu, C., Li, C.Y. & Kong, A.N. Induction of phase I, II and III drug metabolism/transport by xenobiotics. *Arch Pharm Res* 28, 249-268 (2005).

89. Yan, H., Yuan, W., Velculescu, V.E., Vogelstein, B. & Kinzler, K.W. Allelic variation in human gene expression. *Science (New York, N.Y* 297, 1143 (2002).

90. Yoon, S., Yang, Y., Choi, J. & Seong, J. Large scale data mining approach for gene-specific standardization of microarray gene expression data. *Bioinformatics* 22, 2898-2904 (2006).

Samenvatting

Inleiding

Met de dagelijkse inname van voedsel komen er naast voedingsstoffen ook schadelijke stoffen binnen zoals *pathogenen* (bacteriën en virussen). Deze schadelijke stoffen komen terecht in de darm. De wand van de darm is het grootste grensoppervlak tussen de buitenwereld en het lichaam. In een volwassen persoon is het oppervlak bijvoorbeeld 100 keer groter dan het huidoppervlak. De darmwand is uitgerust met elementen die voorkomen dat schadelijke stoffen in het lichaam door kunnen dringen en zo ziekten veroorzaken. Ten eerste zijn er darmcellen die het oppervlak bekleden met een nauw gesloten laag. Deze darmcellen scheiden antibacteriële moleculen uit en ze maken een slijmlaag die het darmoppervlak bedekt. De slijmlaag vormt een barrière tegen pathogenen en beschermt de darmcellen tegen schadelijke objecten. Een tweede barrière bestaat uit een groot aantal darmflorabacteriën. Deze "goede" bacteriën zorgen er voor dat de slechte bacteriën geen voedsel kunnen vinden en dat ze zich niet aan de darmwand kunnen binden. Ten derde is er het immuunsysteem dat ongewilde indringers uitschakelt. Deze verschillende elementen beïnvloeden elkaar en vormen samen een effectieve barrière tegen pathogenen en andere schadelijke stoffen.

Toch komen darminfecties, veroorzaakt door bacteriën, virussen of andere schadelijke stoffen, nog regelmatig voor. Vooral ouderen, zieken en baby's zijn extra gevoelig vanwege hun zwakkere afweersysteem. Versterking van de darmbarrière kan zorgen voor een betere bescherming. Er staat tegenwoordig een groeiend aantal producten in de supermarkt die claimen de darmgezondheid te bevorderen en zo de weerstand te verhogen. Wat deze stoffen precies doen en of ze echt een verbetering van de barrière geven is vaak onduidelijk.

Doel van dit proefschrift is het identificeren van biologische processen die van belang zijn voor de barrièrefunctie van darmcellen in een levend dier, zodat we beter begrijpen hoe darminfecties en verstoring van darmfuncties (zoals verhoogde darmdoorlaatbaarheid) ontstaan. De uitkomsten moeten bijdragen aan de ontwikkeling van voedingstoffen die de darmgezondheid verbeteren. Vooral de vroege reacties van de darmwand op pathogenen en schadelijke stoffen zijn interessant, omdat in deze fase erger kan worden voorkomen.

Aanpak

Om de darmbarrière beter te begrijpen, is het nodig om effecten van pathogenen en voedingsmiddelen op de darm in een levend organisme te bestuderen. Tot nu toe is onderzoek naar de darmbarrière vooral gedaan door middel van zogenaamde *in vitro-* modellen. Dit zijn celkweekstudies of bacteriecultures, waarvan de verschillende elementen in een reageerbuis of op een petrischaal apart onderzocht worden. De onderlinge samenhang tussen de verschillende elementen gaat dan verloren. Dat maakt het lastig de resultaten te vertalen naar mens of dier. Deze vertaalslag is echter noodzakelijk om de resultaten van het onderzoek, naar effecten van voeding of medicijnen op de darmbarrière, uiteindelijk te kunnen toepassen voor het voorkomen van darmontstekingen bij mensen. In dit onderzoek hebben we daarom gebruik gemaakt van levende ratten als proefdieren.

Om de biologische processen in de darm van de rat te kunnen identificeren maakten we gebruik van *transcriptomics*. Dit is een methode die het mogelijk maakt om van een groot aantal genen tegelijk de *genexpressie* te meten. De genexpressie is een afspiegeling van de biologische processen die actief zijn in het onderzochte weefsel, in dit geval de binnenste laag van de darmwand van een rattendarm. We vergeleken de genexpressie-profielen (een verzameling datapunten die aangeven welk gen er actief is en welke niet) in de darm van behandelde ratten met onbehandelde ratten. Zo kregen we inzicht in alle biologische processen, die op een bepaald tijdstip verschilden tussen de behandelde groep en de onbehandelde (controle)groep.

We hebben de darmwand geprikkeld met schadelijke stoffen en het effect van deze prikkeling op de genexpressie in de darmwand gemeten. We hebben twee verschillende stoffen gebruikt die stress van de darmbarrière veroorzaken: ten eerste het veel voorkomende voedselpathogeen *Salmonella* en ten tweede het voedingsmiddel *Fructo-oligosacchariden (FOS)*.

Salmonella

Salmonella enteritidis is wereldwijd een van de meest voorkomende oorzaken van voedselinfecties. Het veroorzaakt ontsteking van de darmwand en diarree. Het verloop van een Salmonella-infectie bij de mens komt overeen met het verloop van een infectie bij ratten. De ratten werden geïnfecteerd met Salmonella en op drie verschillende momenten na de infectie hebben we darmweefsel van het *ileum* (dunne darm) en het *colon* (dikke darm) verwijderd van geïnfecteerde en van controle dieren. Vervolgens hebben we de genexpressie van dit weefsel geanalyseerd. De resultaten van deze analyses in het ileum staan beschreven in hoofdstuk 2 en de resultaten van het colon in hoofdstuk 3.

FOS

Er is veel wetenschappelijke en commerciële belangstelling naar *prebiotica*. Dit zijn niet-verteerbare voedingsvezels, zoals FOS, die in de dunne darm niet afgebroken worden door de spijsverteringsenzymen en daardoor onverteerd de dikke darm bereiken. Daar vormen ze een voedingsbron voor de darmflora (*fermentatie*). Er is wetenschappelijk bewijs dat FOS daardoor de groei van 'goede' darmbacteriën stimuleren. Deze toename in darmflorabacteriën wordt gezien als gezond en vormt de wetenschappelijke basis waarop prebiotica aan zuivelproducten, zuigelingenvoeding en brood worden toegevoegd. Voorgaande studies naar het effect van FOS op de darmen van ratten hebben - in tegenstelling tot de algemene verwachting - aangetoond dat een FOS-dieet een negatieve invloed heeft op de darmgezondheid. Ratten op een FOS-dieet hebben een hogere darmdoorlaatbaarheid en een verhoogde passage van de Salmonellabacterie door de darmwand dan ratten op een controledieet. Verhoogde darmdoorlaatbaarheid is een teken van beschadiging van de darmbarrière en verhoogt de kans dat schadelijke stoffen en bacteriën de darmbarrière doordringen. Door middel van welk mechanisme FOS de darmbarrière verzwakken is nog niet bekend. Om deze mechanismen te bestuderen hebben we de reactie van de darm op FOS en Salmonella onderzocht in ratten. We hebben ratten twee weken een dieet gegeven met FOS of een controledieet. Een deel van de dieren hebben we geïnfecteerd met Salmonella. Aangezien FOS werkzaam zijn in het colon, hebben we ons onderzoek daarop gericht. De resultaten van het effect van FOS op de verhoogde passage van de Salmonella bacterie door de darmwand staan beschreven in hoofdstuk 3. De resultaten van het effect van FOS op de darmdoorlaatbaarheid staan in hoofdstuk 6.

Resultaten

Barrière mechanismen

Met de genexpressie-analyse na infectie met Salmonella en FOS hebben we nieuwe biologische processen aangetoond die een rol spelen in de darmbarrière. Deze processen zijn niet eerder betrokken in onderzoek naar weerstand tegen een pathogeen of een andere schadelijke stof. Een van de processen die we gevonden hebben, is bijvoorbeeld dat FOS veel genen veranderen die voor de *energie-huishouding* in de cel zorgen. Het is bekend dat een verstoorde energie-huishouding van darmcellen verhoogde darmdoorlaatbaarheid kan veroorzaken. Nieuw is onze bevinding dat FOS door het mechanisme van verstoorde energie-huishouding de darmdoorlaatbaarheid verhogen.

De genexpressie-analyses tonen verder aan dat Salmonella en FOS de genen, waarvan we verwachten dat ze verantwoordelijk zijn voor de barrièremechanismen, niet beïnvloeden. Deze genen zijn bijvoorbeeld de genen verantwoordelijk voor de nauwe aansluiting van de darmcellen en de genen verantwoordelijk voor de slijmlaagproductie. Daarnaast vonden we dat Salmonella in de rattendarm minder genen beïnvloedt dan in *in vitro*-modellen.

De aanpak van ons darmbarrière-onderzoek in levende dieren benadrukt dat:

1) extrapolatie van resultaten uit *in vitro*-modellen naar de situatie in een levend organisme, in het geval van darmbarrière-functie, niet altijd relevant is en dus met grote voorzichtigheid moet gebeuren;

2) darmbarrière-onderzoek zich niet alleen moet richten op verwachte barrièreprocessen: andere onverwachte genen en processen kunnen belangrijk zijn in de vroege response op een pathogeen of schadelijke stof.

FOS en Salmonella

In de ratten op het FOS dieet veranderden meer genen door Salmonella dan in ratten op een controledieet. In de FOS dieet groep vonden we onder andere een verhoging van het aantal immuungenen, wat in prebiotica onderzoek ook wel gebruikt wordt om aan te tonen dat prebiotica een verhoogde bescherming biedt. Onze studie heeft aangetoond dat een verhoging van het aantal immuungenen in de darm gerelateerd is aan een verhoogde passage van Salmonella door de darmwand. En dus dat een verhoging van het aantal immuungenen een reactie is op stress van de darmwand en niet een verhoogde weerstand betekent. Dit betekent dat verhoging van immuungenen voorzichtig geïnterpreteerd moet worden, bij voorkeur altijd gekoppeld aan meetbare eindpunten zoals passage van de bacteriën door de darmwand.

Plaats in de darm

We hebben aangetoond dat Salmonella-infectie in ratten zeker niet beperkt is tot het ileum, zoals algemeen wordt aangenomen. Een Salmonella-infectie veroorzaakte in een vroeg stadium namelijk een vergelijkbare genexpressie-reactie in het colon en het ileum.

Functionele studies naar de mogelijke biomarker PAP

Biomarkers zijn biologische indicatoren die gebruikt kunnen worden om ziekten aan te tonen, het verloop van ziekten te volgen en om effecten van de omgeving op ziekten te meten. In voedingsstudies maken biomarkers het mogelijk om bijvoorbeeld de werking van een nutriënt (voedingsmiddel) op darmgezondheid te meten. In het darmgericht voedingsonderzoek is behoefte aan nieuwe biomarkers die de conditie van de darmbarrière kunnen meten en volgen.

We hebben een aantal genen gevonden die mogelijk geschikt zijn als biomarker. Van één gen, PAP, hebben we in hoofdstuk 4 de bruikbaarheid als biomarker nader bestudeerd door het eiwitproduct te onderzoeken. Voordat PAP als nieuwe biomarker voor darmgezondheid geaccepteerd kan worden, zijn echter aanvullende studies nodig, bijvoorbeeld om uit te zoeken of het eiwit zich in de mens hetzelfde gedraagt.

Data analyse in transcriptomic studies

De gegevens uit transcriptomic (genexpressie) studies zijn zeer complex door de enorme hoeveelheid informatie (het hoge aantal genen). Daar komt bij dat de effecten op genexpressie in voedingsstudies klein zijn en er grote variatie in genexpressie binnen groepen gevonden wordt. Dit was ook het geval in onze Salmonella en FOS studies. Deze combinatie van veel genen, kleine effecten en grote variatie maakt het ingewikkeld om te bepalen welke genen betrouwbaar veranderd zijn en welke genen bij toeval veranderd lijken, ofwel tot de ruis behoren. Gebruikelijke statistische tests bleken te strikt voor de kleine effecten veroorzaakt door Salmonella en FOS. In hoofdstuk 5 bestudeerden we een nieuwe statistische methode die geschikt is om kleine genexpressie-verschillen te analyseren. Met deze methode waren we in staat om een objectieve scheidslijn te bepalen tussen genen die significant veranderd waren en genen die tot de ruis behoorden. Deze methode kan in de toekomst nuttig zijn in humane genexpressie-studies, waar individuele variatie nog groter is dan in de studies met ratten.

Conclusies, toekomst

De studies in dit proefschrift laten zien dat het gebruik van transcriptomics (genexpressie-studies) een succesvolle aanpak is voor het identificeren van nieuwe biomarkers (zoals PAP) en biologische processen (zoals energie-huishouding), die belangrijk zijn bij de barrièrefunctie van de darm in ratten.

De studies tonen aan dat onderzoek naar de darmbarrière niet alleen gericht moet zijn op een kleine groep "bekende" barrière genen. Daarnaast is voorzichtigheid geboden bij het extrapoleren van data uit *in vitro*-modellen naar levende organismen en bij het positief interpreteren van verhoogde weerstand markers zonder koppeling met meetbare eindpunten.

De volgende stap in darmbarrière-onderzoek is om de in dit onderzoek aangetoonde resultaten op het gebied van genexpressie te bevestigen in vervolgonderzoek, zodat ze uiteindelijk gebruikt kunnen worden om wetenschappelijk aan te tonen of een voedingstof of medicijn de darmgezondheid daadwerkelijk beïnvloedt.

Supplemental data

Color Figure Chapter 1

BOX 1. Transcriptomic technique

The sequences of thousands of genes are available for whole genomes; this allows the design of sequence fragments or probes representing individual genes. These gene-specific probes are printed on nylon membranes, glass slides or silicon chips. Hybridization of single stranded DNA allows determination of mRNA levels of all genes spotted on the array.

Procedure: 1) Isolation of mRNA from biological samples. 2) Fluorescent label incorporation into cDNA copy of the mRNA. 3) Hybridization of labeled sample to the array: each cDNA anneals to the complementary cDNA probe for a specific gene on the surface. 4) Scanning the array to determine the fluorescent amount of labeled cDNA hybridized to each spot. 5) Quantification of the fluorescent signal, and data normalization to enable comparison between the different arrays.

Spot ID	Gene ID	Intensity signal			
		Array 1	Array 2	Array 3	Array 4
1	NM_031752	8.318	7.790	8.343	8.041
2	AW918040	9.458	9.533	9.397	9.563
3	XM_340892	8.196	7.858	8.295	7.969
4	XM_213821	9.742	9.455	9.376	9.304
5	XM_218624	9.090	8.655	8.489	8.975
6	XM_341784	8.071	8.261	7.990	8.287
7	XM_232246	8.469	7.995	8.023	8.488
8	NM_012527	8.616	8.812	8.032	7.680
9	NM_080781	8.469	8.085	7.917	8.232
10	AW918123	10.301	9.862	9.381	9.992
11	XM_341315	8.048	7.096	7.618	7.580
12	XM_214047	10.804	11.278	10.532	9.905
13	BG665384	10.085	9.905	9.475	9.990

Color Figure Chapter 2

A

B

Figure 1. The number of differentially expressed genes with a fold change greater than 2 in ileal mucosa and ileal Peyer's patches of rats orally infected with Salmonella at 1, 3 or 6 days postinfection compared to sham treated controls (A).

Hierarchical clustering of genes with a fold change greater than 2 (infected/control). Genes and experiments were clustered using Pearson UPGMA (GeneMathsXT). Color scale represents ratio infected/control (B).

Supplemental data Chapter 2

Supplemental table 2.1 Sequences of the primers used for Q-PCR analysis.

Gene symbol	Sequence ID	Forward primer (5' 3')	Reverse primer (5' 3')	Product length
Actb	NM_031144	CTTTCTACAATGAGCTGCGTGTG	GTCAGGATCTTCATGAGGTAGTCTGTC	315
Cyp1a1	NM_012540	CCCACAGCACCATAAGAGATACAAG	GCCCAAACCAAAGAGAATGACCTTC	200
Cyp3a9	U60085	CTGTATTGGCATGAGGTTTGCTCTC	AGAAGCAGTGGCTTTTCTGGTTG	150
Cyp4f14	NM_019623	CTCACCCCTGATGGGATGCG	TCCGACAGCTCCTTTCCATCTTC	200
Rd5	XM_214386	TGTCCTCCTTTCTGCCCTTGTC	TCAGCGGCAACAGAGTATGAGG	271
Ephx1	M26125	CCTTCTGGGCTTTGTCATCTACTGG	CCTCTGGTGTAAGTCCTTGATCTCC	172
Gro	D11444	ACCGAAGTCATAGCCACACTCAAG	CACCAGACAGACGCCATCGG	150
Lcn2	NM_130741	TCTCTGGGCCTCAAGGATAACAAC	AGGAAAGATGGAGCGGCAGAC	150
Lys	NM_012771	CAATGTGCGAAGAGAGTTGTGAGG	AGAGACAGTGTGAGCTGAGTAGAAG	149
Mmp7	NM_012864	CGACATTGCAGGCATCCAGAAG	GGAGTAAGTGTGGCTCAGGAAGG	105
NFκB	L26267	AAGTGATCCAGGCAGCCTTCC	TTCAGAGATAGCAGTGGGCCATC	220
Pap	L07127	CTGCCAGAAGAGACCTGAAGGAC	CACCTCCATTGGGTTCTCCACC	154
PapIII	U09193	GCTTCCTTTGTGTCCTCCTTGATT	TACTCCACTCCCATCCACCTCTG	116
Plekha3	NM_001013077	GCAGAACCTACTCGGACACAGAC	TCTTCCCGAATGGCTGCTGATG	150
Slc2a2	NM_012879	ACCAACTGGCTCTTGTCACAGG	GGTACAGCAGATAGGCCAAGTAGG	109
Slc2a5	NM_031741	CCCTGACTGTCTCCATGTTCCC	CCCTTTTCTGCCCAAGTTATTCACC	86
Slc5a1	NM_013033	TTTATCCTGACTGGGTTTGCTTTCC	GTCGGCCCTGGGTGTGTAAC	138
Timp1	NM_053819	TGGTTCCCTGGCATAATCTGAGC	CACAAGCAATGACTGTCACTCTCC	136
Tlr5	AY197552	GATTTGTTTCCTGTGTTTTAAGACCATCC	TGAAGCAGAAGTAGGCGTCGTATC	104
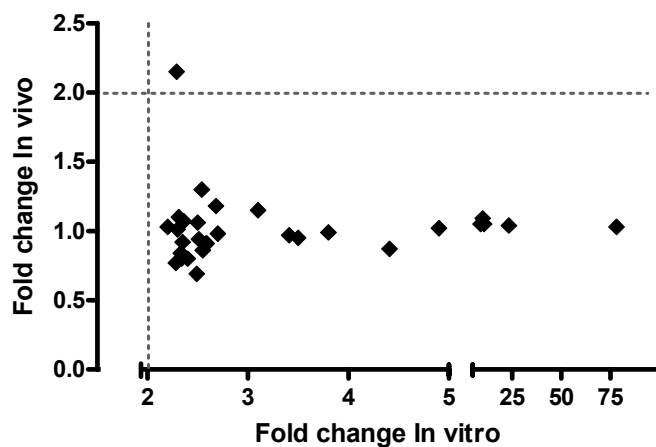
Supplemental figure 2.1 Salmonella affected genes in vitro vs. in vivo.

Scatter plot for 30 genes that were more than twofold altered *in vitro* in HT-29 cells 3 hours after Salmonella infection (Eckmann *et al*) compared to the *in vivo* fold expression changes of the same genes in ileal mucosa 1 day after Salmonella infection. No correlation was found (r^2=0.0004).

Supplemental table 2.2 Salmonella affected genes *in vitro* vs *in vivo.*

Gene name	Accession number		*In vitro*	*In vivo*
	Human	Rat	3 hours p.i.	1 day p.i.
Macrophage inflammatory protein-2	X53799	NM_053647	78.0	1.0
Tyrosine kinase related to TRK/ Ddr2	X74764	NM_031764	23.4	1.0
Interferon regulatory factor-1	X14454	NM_012591	10.8	1.0
Inhibinβ A-subunit	J03634	NM_017128	10.1	1.1
G-CSF	X03438	NM_017104	9.1	1.0
Glial growth factor	L12261	U02316	4.9	1.0
Cysteine protease Mch2α/ Caspase 6	W45688	NM_031775	4.4	-1.2
Platelet-derived growth factor, beta chain	X02811	AF359356	3.8	-1.0

Follistatin-related protein	U06863	NM_024369	3.5	-1.1
Cysteine-rich fibroblast growth factor receptor 1/ Glg1	H66617	NM_017211	3.4	-1.0
CC chemokine receptor 6	N57964	NM_019310	3.1	1.2
CXCR4	D10924	NM_022205	2.7	-1.0
LI-cadherin	AA088861	NM_053977	2.7	1.2
Ubiquitin-conjugating enzyme E2	H20743	NM_013050	2.6	-1.1
Glucosidase I	AA291490	NM_031749	2.6	-1.1
Nuclear-specific cyclophilin-60	AA682506	NM_017101	2.6	-1.2
Interleukin 8	AA102526	NM_138522	2.5	1.3
Ribosomal protein L27a	AA599178	X52733	2.5	-1.1
FMLP receptor homologue-1	M76673	NM_022218	2.5	1.1
Transfer RNA synthetase-like protein	W96450	Rn.8968	2.5	-1.5
Macrophage migratory inhibitory factor	M25639	NM_031051	2.4	-1.3
BGT-1	N49856	NM_017335	2.4	1.1
Heterogeneous nuclear ribonucleoprotein	AA490991	Rn.23677	2.4	-1.1
Acidic calponin 3	AA043228	NM_019359	2.3	-1.3
HLA Class I, C heavy chain	AA464246	AF074609	2.3	-1.2
CCAAT displacement protein	AA292536	Rn.10455	2.3	1.1
BCNT craniofacial development protein 1	AA682613	Rn.128746	2.3	1.0
HLA Class II	AA669055	X56596	2.3	2.1
Guanosine 5'-monophosphate synthase	N59764	Rn.12593	2.3	-1.3
Insulin receptor	M10051	NM_017071	2.2	1.0

Supplemental table 2.3 Salmonella induced gene expression changes in ileal mucosa.
Genes with a fold-change in expression greater than 2.0 are given for day 1, 3 and 6 postinfection in ileal mucosa between rats that were orally infected by *Salmonella* relative to their sham-treated controls. The file is available at the website of Physiological Genomics (http://physiolgenomics.physiology.org/).

Supplemental table 2.4 *Salmonella* induced gene expression changes in Peyer's patches.
Genes with a fold-change in expression greater than 2.0 are given for day 1, 3 and 6 postinfection in Peyer's patches between rats that were orally infected by *Salmonella* relative to their sham-treated controls. The file is available at the website of Physiological Genomics (http://physiolgenomics.physiology.org/).

Supplemental data Chapter 3

Additional File 3.1. Salmonella affected colonic genes.
Genes that were upregulated of downregulated at least 1.5-fold in rat colon mucosa by Salmonella at days 1, 3, and 6 after oral Salmonella infection compared to colon mucosa of non-infected rats. This file is available at the BMC website (http://www.biomedcentral.com).

Supplemental table 3.1 Primer sequences of the primers used for Q-PCR analysis.

Gene symbol	Sequence ID	Forward primer (5' 3')	Reverse primer (5' 3')	product length
Actin	NM_031144	CTTTCTACAATGAGCTGCGTGTG	GTCAGGATCTTCATGAGGTAGTCTGTC	315
Arf1	NM_022518	CTCCGAGATGCCGTTCTCTTG	GGTAGCCTGAATGTACCAGTTCC	128
Clca6	NM_201419	GATGGGGAAAAGATCAGTCTAACATGG	GTTTCTGGTTTAAAGGCAAAGGTTTCC	200
Gbp2	XM_225909	GACCTCAAGCCTAGAGCACAC	GACTTCAAGCAAATAAAGCCACAG	106
Gpx2	NM_183403	GCCTAGTGGTTCTCGGCTTCC	AGGGTAGGGCAGCTTGTCTTTC	200
Ifi47	NM_172019	GTGCGGTTGGTGGTGGTTG	CCGAGTCTGTTGCTCACTGC	83
Il1b	NM_031512	AAGGGGGTTGAATCTATACCTGTCCTG	TGCTCTGCTTGAGAGGTGCTG	200
Pap	NM_053289	CTGCCAGAAGAGACCTGAAGGAC	CACCTCCATTGGGTTCTCCACC	154
Pla2g2a	NM_031598	CAAAGTTTCTGACCTACAAGTTCTCCTAC	CTTTCAGCAACTGGGCGTCTTC	200
Rps29	NM_012876	CCGACAGTGCTTCCGTCAG	GACAGTTGGTTTCATTGGGTAGAC	102
Stat1	NM_032612	GTTCGCCACCATCCGCTTC	TCTTCCTCTCCTCCTTCAGACAG	200
Tgm2	NM_019386	CACTTTCTGATTCCCTGTATGACTGTG	ACCCTTGACCGACTTCAGCTTG	200

Supplemental table 3.2 Housekeeping and cell type specific genes.
The fold change in expression of housekeeping genes and cell-type specific genes in the colon mucosa at days 1, 3 and 6 after oral *Salmonella* infection.

Gene Name	Gene symbol	Sequence ID	Fold Change infected vs non-infected rats on different days p.i.			Dietary infection study[a]
			Time course infection study			
			Day 1	Day 3	Day 6	Day 2
Housekeeping genes						
Actin beta	Actb	NM_031144	1.1	1.2	1.1	1.2
Glyceraldehyde-3-phosphate dehydrogenase	Gapdh	NM_017008	1.0	1.0	1.0	1.0
Aldolase a, fructose-biphosphate	Aldoa	NM_01495	1.1	1.0	1.0	-1.1
Phosphoglucerate kinase 1	Pgk	NM_053291	1.0	1.2	1.2	1.1
Adp-ribosylation factor 1	Arf	NM_022518	1.0	1.0	1.0	1.1
H3 histone, family 3B	H3f3b	NM_053985	1.0	-1.1	1.0	1.0
Succinate dehydrogenase complex, subunit A	Sdha	NM_0130428	1.0	-1.1	-1.1	1.0
Ribosomal protein S26	Rps26	NM_013224	-1.1	-1.1	1.0	1.1
Pleckstrin homology domain-containing family A member 3	Plekha3	NM_001013077	1.0	-1.1	1.0	1.1
Cell type specific genes						
Enterocyte						
Intestinal fatty acid binding protein 2	Fabp2	NM_013068	-1.6	-1.5	-2.0	-1.5
Villin 2	Vil2	NM_019357	1.2	1.1	1.0	1.0
Intestinal alkaline phosphatase-II gene	Alpi2	NM_022680	1.1	1.0	-1.0	-1.3
Goblet Cell						
Colonic mucin 2	Muc2	TC556623	-1.1	-1.0	-1.1	-1.1
Mucin 3	Muc3	U76551	1.7	1.3	1.3	1.0
Trefoil factor 1	Tff1	NM_057129	-1.0	-1.0	-1.2	-1.2
Trefoil factor 3	Tff3	NM_013042	1.1	-1.4	1.0	1.2
Chloride channel calcium activated 3	Clca3	XM_217689	1.1	-1.5	-1.2	-1.2

Gene Name	Gene symbol	Sequence ID			
Chloride channel calcium activated 6	Clca6	NM_201419	2.3	2.2	3.7	2.3
Sialyltransferase 8 B	Siat8b	NM_057156	-1.0	1.0	1.0	ND
Sialyltransferase 9	Siat9	NM_031337	1.1	1.0	1.0	1.2
Leukocytes						
Leukocyte cell derived chemotaxin 1	Lect1	NM_030854	1.2	1.2	1.1	1.1
CD84 leukocyte antigen	CD84	XM_577290	1.1	-1.0	1.5	ND
SP140 nuclear body protein	Sp140	XM_237361	1.1	-1.0	1.2	1.1
Leucocyte specific transcript 1	Lst1	NM_022634	-1.0	1.2	1.1	-1.0
RT1 class I, CE12	RT1-CE12	XM_227986	1.2	1.0	1.5	1.1
RT1 class Ib, locus Aw2	RT1-Aw2	Y13890	1.1	-1.0	1.6	1.0
Granzyme A	Gzma	NM_153468	1.0	-1.0	1.7	1.0
Granzyme B	Gzmb	NM_138517	-1.0	-1.0	1.1	1.1
Granzyme C	Gzmc	NM_134332	1.1	1.0	1.3	-1.1

[a] Fold Change infected vs non-infected rats fed a cellulose diet at day 2 p.i (obtained from the dietary infection study).

Genes not detected in this independent study are indicated by ND.

Supplemental table 3.3 Dietary modulated genes.

Processes regulated in colon by *Salmonella* at day 2 in cellulose fed and FOS fed rats.

Gene Name	Gene symbol	Sequence ID	Fold Change infected vs non-infected rats[a]	
			Cellulose diet	FOS diet
Overlap (17 genes)				
Transport				
Chloride channel calcium activated 6	*Clca6*	NM_201419	**2.3**	**3.9**
Oxidative stress				
Glutathione peroxidase 2	*Gpx2*	NM_183403	**2.2**	**4.2**
Immune response, Antimicrobial defense and Inflammatory response				
Pancreatitis-associated protein	*Pap*	NM_053289	**17.7**	**35.6**
Phospholipase A2, group IIA (platelets, synovial fluid)	*Pla2g2a*	NM_031598	**7.3**	**18.9**
Interleukin 1 beta	*Il1b*	NM_031512	**2.6**	**3.5**
TRAF2 binding protein	*Traf2bp*	NM_001014044	**2.4**	**3.4**
Chemokine (C-C motif) ligand 2	*Ccl2*	NM_031530	**2.8**	**3.2**
Chemokine (C-X-C motif) ligand 10	*Cxcl10*	NM_139089	**2.3**	**3.2**
Interferon				
Interferon gamma inducible protein	*Ifi47*	NM_172019	**2.6**	**5.2**
Interferon-inducible GTPase	*Iigp1*	NM_001012353	**2.4**	**3.4**
Proteolysis				
Ubiquitin D	*Ubd*	NM_053299	**3.4**	**7.6**
Proteosome (prosome, macropain) subunit, beta type 9	*Psmb9*	NM_012708	**2.0**	**3.1**
Miscellaneous				
Unknown	-	BF555121	**3.9**	**10.4**
Unknown	-	TC555318	**2.4**	**4.2**
Unknown	-	AI234967	**2.7**	**3.9**
Palmitoyl-protein thioesterase	*Ppt*	XM_342904	**2.8**	**3.9**
Neurexin 2	*Nrxn2*	NM_053846	**-3.3**	**-3.1**
Cellulose only (3 genes)				
Protease, serine, 3 (mesotrypsin)	*Prss3*	XM_342668	**10.7**	-1.7
Unknown	-	TC526384	**2.2**	1.5
Unknown	-	NM_001014241	**-4.4**	-1.4
FOS only (58 genes)				
Transport				
Solute carrier family 10, member 2	*Slc10a2*	NM_017222	1.2	**2.1**
Oxidative stress				

Dual oxidase 2	*Duox2*	NM_024141	1.6	**2.9**
Xanthine dehydrogenase	*Xdh*	NM_017154	1.4	**2.7**
Immune response, Antimicrobial defense and Inflammatory response				
Myxovirus (influenza virus) resistance 2	*Mx2*	NM_134350	1.4	**3.0**
Schlafen 3	*Slfn3*	NM_053687	2.0	**3.0**
Interleukin 1 beta	*Il1b*	NM_031512	1.9	**2.6**
Serum amyloid A 3	*Saa3*	AY325259	1.4	**2.5**
Regenerating islet-derived 3 gamma	*Reg3g*	NM_173097	1.9	**2.4**
Fatty acid binding protein 5, epidermal	*Fabp5*	NM_145878	1.6	**2.3**
Lymphocyte antigen 6 complex, locus E	*Ly6e*	NM_001017467	1.5	**2.3**
Chemokine (C-C motif) ligand 7	*Ccl7*	NM_001007612	2.0	**2.2**
Lipopolysaccharide binding protein	*Lbp*	NM_017208	1.3	**2.0**
Interleukin 1 alpha	*Il1a*	NM_017019	1.4	**2.0**
Toll-like receptor 2	*Tlr2*	NM_198769	1.3	**2.0**
Interferon				
Interferon-inducible GTPase	*Iigp*	NM_001024884	1.7	**3.7**
guanylate nucleotide binding protein 2	*Gbp2*	NM_133624	1.7	**3.4**
signal transducer and activator of transcription 1	*Stat1*	NM_032612	1.8	**3.0**
interferon, alpha-inducible protein	*G1p2*	XM_216605	1.5	**2.6**
indoleamine 2,3-dioxygenase	*Indo*	NM_023973	1.5	**2.3**
interferon gamma induced GTPase	*Igtp*	XM_220451	1.3	**2.3**
interferon inducible protein 1	*Ifi1*	NM_001012007	1.4	**2.3**
Interferon regulatory factor 7	*Irf7*	NM_001033691	1.3	**2.2**
EF hand domain containing 2	*Efhd2*	NM_001031648	1.3	**2.1**
Proteolysis				
transglutaminase 2, C polypeptide	*Tgm2*	NM_019386	1.8	**4.4**
proteosome (prosome, macropain) subunit, beta type 8	*Psmb8*	NM_080767	1.7	**2.7**
protease, serine, 22	*Prss22*	XM_220222	1.6	**2.4**
serine (or cysteine) peptidase inhibitor, clade G, member 1	*Serping1*	NM_199093	1.7	**2.4**
potential ubiquitin ligase	*Herc6*	XM_342700	1.5	**2.4**
carboxypeptidase B gene, exons 6, 7, and 8	*Carb7*	M23953	1.5	**2.2**
stefin A2	*Stfa2*	NM_001004129	1.2	**2.2**
Parkin	*Park2*	NM_020093	1.4	**2.1**
Legumain	*Lgmn*	NM_022226	1.4	**2.1**
stefin 2-like	*Stf2*	XM_221409	1.2	**2.1**
proteasome (prosome, macropain) subunit, beta type 10	*Psmb10*	NM_001025637	1.5	**2.0**
Miscellaneous				
Unknown	-	XM_225905	1.6	**3.6**
adipocyte complement related protein/ Adiponectin	*Acdc*	NM_144744	1.3	**2.9**
Unknown	-	XM_221401	1.5	**2.6**
tripartite motif protein 15	*Trim15*	XM_227945	1.6	**2.4**
matrix Gla protein	*Mgp*	NM_012862	1.7	**2.4**
glucosaminyl (N-acetyl) transferase 3, mucin type	*Gcnt3*	NM_173312	1.8	**2.4**
membrane-spanning 4-domains, subfamily A, member 12	*Ms4a12*	XM_219588	1.4	**2.4**
receptor-interacting serine-threonine kinase 3	*Ripk3*	NM_139342	1.4	**2.4**
Unknown	-	XM_221376	1.4	**2.3**
Unknown		CO402999	1.4	**2.3**
zymogen granule protein 16	*Zg16*	NM_134409	1.2	**2.2**
histidyl tRNA synthetase 2	*Hars2*	BQ207049	1.2	**2.2**
Unknown	-	AA924770	1.1	**2.1**
round spermatids protein STDP2	*Stdp2*	XM_573991	1.2	**2.1**
Unknown	-	AW920888	1.3	**2.1**
B-cell leukemia/lymphoma 2 related protein A1	*Bcl2a1*	NM_133416	1.7	**2.1**
GLI pathogenesis-related 2	*Glipr2*	XM_342827	1.4	**2.1**

Unknown	-	XM_223906	1.2	**2.1**
Unknown	-	XM_575189	1.4	**2.1**
Unknown	-	NM_001014209	1.5	**2.0**
MAp19 protein	-	Y18568	1.2	**2.0**
basic transcription element binding protein 1	*Bteb1*	NM_057211	1.3	**2.0**
Unknown	-	BF281337	1.2	**2.0**
period homolog 2 (Drosophila)	*Per2*	NM_031678	-1.1	**-2.2**

[a] Values in bold exceed cut-off value FC>2 or FC<-2.

Color figure Chapter 4

A

12.5x

B

40x

| non-infected | control | calcium |

infected

Figure 4. Pancreatitis associated protein/regenerating gene III (PAP/RegIII) protein expression in the ileal mucosa of non-infected and Salmonella infected rats fed either the control or calcium-supplemented diet. PAP/RegIII protein was present in epithelial cells at the crypt-villus junction (A, 12.5x; B, 40x). Immunostaining for lysozyme precisely identified paneth cells at the base of the crypts were PAP/RegIII staining was absent (data not shown). The number of PAP/RegIII positive cells increased after infection, especially in the direction of villus tips (A). Dietary calcium reduced PAP/RegIII expression in infected animals. No staining was observed in sections incubated without primary antibody (data not shown). All sections were counterstained with haematoxylin to visualise nuclei.

Color figures Chapter 5

Figure 2 a and **b**. Genes, of 100 random sets (black lines) and real sets with different seed values (colored lines), ranked by the I_m values. For colon (A) and cecum (B) datasets.

BOX 1. A framework for identification of physiological responses in microarray based gene expression studies. The framework is composed of the following steps: Gene ranking, gene selection, gene grouping and biological interpretation. Essential features of the data-analysis framework are that 1) all genes (annotated and non-annotated) in the dataset are included in the analysis, 2) interaction between genes is taken into account and, 3) an objective threshold is used for selection of a well-defined gene set. Random forest has these features. Gene grouping can provide information on new targets and add information above pathway analysis. Despite loss of information due to incomplete annotation of the complete dataset, Gene set enrichment analysis can provide additional information on related genes with small differences.

Supplemental data Chapter 6

Supplemental table 6.1 Primer sequences of the primers used for Q-PCR analysis.

Gene symbol	Acc. Nr	Sense	Anti-sense	Product length
ActB	NM_031144	CTTTCTACAATGAGCTGCGTGTG	GTCAGGATCTTCATGAGGTAGTCTGTC	315
Akr1b8	NM_173136	TTCACTATTACAGGACCCCAAGATTAAAG	CGCTCAACTGGAAGTCAAAGACC	170
Atp5i	XM_001075306	TGACTTACTCAGAGCCTCGATTAGC	AGCCAACAACACCACGTTTGC	243
Cox7b	NM_182819	AGAAGACACCTACTTTCCATGACAAATATG	TTTAATGAGTACATGATTCTTTGACTTGGC	250
Gcg	NM_012707	AGGGACCTTTACCAGTGATGTGAG	TTCACCAGCCAAGCAATGAATTCC	75
Me1	M30596	CTTTCAGGCGTGACTTACAGTGTAG	CCCAGATAACTACCCTGAGGAAACC	76
Ndufb9	XM_216929	CGAGTGCTACAAGGTTCCAGAATG	TGCGGTGCCTGTCTCTATGTG	273
Nup37	XM_216872	ATCACTCGGTCCAGTTATCCTCAAG	AGCTTATGGTCTCCTCCAACTGC	248
Pla2g2a	NM_031598	TTCTGACCTACAAGTTCTCCTACCG	TCAGCAACTGGGCGTCTTCC	191
Plekha6	XM_341118	TCAATAAGGAGCTATCCACTCCAGAC	GCACAGGAGATTTCAATCCGCTTC	250
Psma31	BN000326	AAGCTGCAAAGACAGAAATAGAAAAGC	TTGTCGTCATCTGATTCATCTTCTTCC	250
Rbp7	XM_575960	CTGGGAGAACGACAAACTCACTTG	CCAATATAAGGCTCTTTATCAACCCAAAAC	250
Mrps16	XM_001064095	CACTACCTAACAGTCATGGAGAAAAGC	CACTGAAGTCAGCTTGCTTCTGTC	250
Sdhb	XM_216558	AGGCTTATCGCTGGATGATCGAC	GAAGGGACTCACGCCAGAGC	250
Tff3	NM_013042	CACATCAGAGCAGTGTAACAACCG	GCTGACTGTAAGGTCTTTATTCTTCTGG	250
Timm8b	NM_022541	TGGTGAAGCGGACGAAGCG	ATGGCAAGAGTAGTGTCAATGAATCG	198
Uqcrb	XM_001074024	TGAGAGATGATACAATACCTGAAACTGAAG	TCTCCTTTCTTTCCCAAATAACCTCTTTC	216

Supplemental table 6.2 The effect of FOS on the expression of potential barrier associated genes.

Gene name	Gene name*	Sequence ID	Fold change†	p-value
Tight junctions				
claudin 2	*Cldn2*	XM_236535	1.0	0.519
claudin 4	*Cldn4*	XM_222088	1.0	0.823
claudin 23	***Cldn23***	XM_224915	1.2	0.001
Occludin	*Ocln*	NM_031329	1.0	0.960
cadherin 1	*Cdh1*	NM_031334	1.0	0.626
cadherin 5	*Cdh5*	XM_226213	1.0	0.924
beta catenin-like 1	*Ctnlb1*	CB546891	1.0	0.575
desmocollin 2	*Dsc2*	XM_226120	-1.2	0.271
junctional adhesion molecule 3	*Jam3*	NM_001004269	1.0	0.955
tight junction protein 1	*Tjp1*	XM_218747	-1.3	0.014
Apoptosis				
apoptosis antagonizing transcription factor	***Aatf***	NM_053720	1.1	0.001
apoptosis related protein p18 protein	***Apr3***	XM_216650	1.3	0.001
apoptotic peptidase activating factor 1	*Apaf1*	NM_023979	1.0	0.637
BAX protein, cytoplasmic isoform delta	***Bax***	AF235993	1.3	<0.001
B-cell CLL/lymphoma 10	*Bcl10*	NM_031328	1.0	0.969
B-cell leukemia/lymphoma 2	*Bcl2*	NM_016993	1.1	0.047
BCL2/adenovirus E1B 19kDa-interacting protein 1	***Bnip1***	NM_080897	1.2	0.001
BCL2-antagonist/killer 1	***Bak1***	NM_053812	1.5	<0.001
Bcl2-associated athanogene 1	***Bag1***	XM_216377	1.2	0.010
bcl2-associated death promoter	*Bad*	NM_022698	1.2	0.012
Bcl2-interacting killer	*Biklk*	NM_053704	1.1	0.299
Bcl2-like 1	*Bcl2l1*	AW142029	1.0	0.948
Bcl-2-related ovarian killer protein	*Bok*	NM_017312	1.1	0.315
BH3 interacting domain death agonist	*Bid*	NM_022684	1.0	0.793
caspase 2	*Casp2*	NM_022522	-1.1	0.269
caspase 3, apoptosis related cysteine protease	*Casp3*	NM_012922	1.2	0.032

caspase 7	*Casp7*	NM_022260	1.2	0.001
catalase	*Cat*	NM_012520	1.3	0.014
CDC42 effector protein (Rho GTPase binding) 5	*Cdc42ep5*	XM_341784	1.2	0.016
CDC42 small effector 1	*Cdc42se1*	AW920756	1.2	0.001
cell division cycle 42 homolog (S. cerevisiae)	*Cdc42*	NM_171994	1.3	0.006
cell division cycle and apoptosis regulator 1	*Ccar1*	XM_342143	1.2	0.001
deoxyribonuclease I	*Dnase1*	NM_013097	2.3	<0.001
programmed cell death 6	*Pdcd6*	XM_217732	1.2	<0.001
programmed cell death 8	*Pdcd8*	NM_031356	1.3	0.001
PYD and CARD domain containing	*Pycard*	NM_172322	1.6	<0.001
STEAP family member 3	*Steap3*	NM_133314	1.4	<0.001
TatD DNase domain containing 1	*Tatdn1*	XM_228158	1.3	<0.001
transmembrane BAX inhibitor motif containing 4	*Tmbim4*	NM_199116	1.2	0.001
v-crk sarcoma virus CT10 oncogene homolog (avian)	*Crk*	BG671506	1.1	0.013
Mucosal defense				
mucin 2	*Muc2*	U07615	-1.2	0.381
mucin 3	*Muc3*	U76551	1.1	0.519
gastric mucin	*Muc*	XM_344685	1.7	<0.001
defensin NP-4 precursor	*Np4*	NM_173299	-1.4	0.026
defensin related cryptdin 4	*Rd5*	XM_214386	-1.1	0.730
defensin, alpha 5, Paneth cell-specific	*Defa*	NM_173329	-1.1	0.090
lipocalin 2	*Lcn2*	NM_130741	1.0	0.635
phospholipase A2, group IIA (platelets, synovial fluid)	*Pla2g2a*	NM_031598	3.7	<0.001
matrix metallopeptidase 7	*Mmp7*	NM_012864	-1.1	0.111
lipopolysaccharide binding protein	*Lbp*	NM_017208	1.7	0.018
cathelicidin antimicrobial peptide	*CRAMP*	CA509601	1.1	0.173
S100 calcium binding protein A8 (calgranulin A)	*S100a8*	NM_053822	1.3	0.091
S100 calcium binding protein A9 (calgranulin B)	*S100a9*	NM_053587	1.4	0.003
toll-like receptor 1	*Tlr1*	XM_223421	1.5	<0.001
toll-like receptor 2	*Tlr2*	NM_198769	1.1	0.294
toll-like receptor 3	*Tlr3*	NM_198791	1.1	0.489
toll-like receptor 4	*Tlr4*	NM_019178	-1.1	0.150
toll-like receptor 5	*Tlr5*	XM_223016	1.1	0.588
toll-like receptor 9	*Tlr9*	NM_198131	1.0	0.732
trefoil factor 1	*Tff1*	NM_057129	1.6	<0.001
trefoil factor 3	*Tff3*	NM_013042	1.7	<0.001
beta-2 microglobulin	*B2m*	NM_012512	1.4	<0.001
MHC class II region expressed gene KE2	*Ke2*	NM_212506	1.4	<0.001
matrix metallopeptidase 7	*Mmp7*	NM_012864	-1.1	0.111
lipopolysaccharide binding protein	*Lbp*	NM_017208	1.7	0.018
immunoglobulin joining chain	*Igj*	XM_341195	1.2	0.422
CD79A antigen (immunoglobulin-associated alpha)	*Iga*	XM_001077003	1.0	0.987
lectin, mannose-binding 2	*Lman2*	XM_214428	1.3	0.001

* Genes significantly regulated by FOS are shown in bold gene symbols.

† Ratio FOS diet/ Control diet.

Supplemental table 6.3 Functional classification of FOS affected genes by Metacore.

	Number of genes		
Process*	Selection	Total	p-value§
Generation of precursor metabolites and energy [1]	53	323	2.3E-17
Translation: Translation in mitochondria [2]	31	187	6.7E-13
Mitochondrial electron transport, NADH to ubiquinone [1]	11	30	2.6E-08
Protein targeting to mitochondrion [1]	10	25	4.3E-08
Mitochondrial transport [1]	12	41	1.1E-07
Electron transport [1]	19	108	1.9E-07

Oxidative phosphorylation [1]	13	52	2.5E-07
ATP synthesis coupled electron transport [1]	11	37	3.1E-07
Proteolysis: Ubiquitin-proteasomal proteolysis [2]	21	167	6.1E-07
Response to hypoxia and oxidative stress [2]	13	93	2.9E-05
Coenzyme metabolic process [1]	14	91	3.8E-05
Regulation of carbohydrate metabolic process [1]	6	17	5.7E-05
Protein thiol-disulfide exchange [1]	3	3	8.4E-05
Immune: antigen presentation [2]	19	196	9.1E-05
Carbohydrate metabolic process [1]	30	328	1.1E-04
Muscle filament sliding [1]	4	8	2.2E-04
Energy derivation by oxidation of organic compounds [1]	16	133	2.3E-04
Cytoskeleton-dependent intracellular transport [1]	11	73	3.1E-04
Cofactor metabolic process [1]	15	125	3.6E-04
Response to copper ion [1]	4	9	3.9E-04
Response to inorganic substance [1]	7	34	5.7E-04
Main pathways of carbohydrate metabolic process [1]	10	68	7.1E-04
Nucleosome assembly [1]	7	36	8.3E-04

 * Biological processes based on [1]Metacore GO-analysis or [2]Metacore GeneGo-analysis.

† The number of genes affected by FOS (selection) and the total number of genes present in the process are given (total).

§ Metacore derived p-value.

Processes containing >500 genes were excluded from the analysis, as these processes represent uninformative broad classes such as "biological function".

Supplemental table 6.4 Genesets enriched in FOS versus control dataset according to GSEA analysis.

Process *	N	ES	NES	q-value†
Oxidative phosphorylation[1]	25	0.65	1.96	0.07
Propanoate metabolism[1]	24	0.62	1.92	0.10
Proteasome[1]	16	0.82	1.83	0.16
Free Radical Induced Apoptosis[2]	8	0.72	1.80	0.19
Butanoate metabolism[1]	20	0.66	1.79	0.19
Programmed cell death[3]	10	0.68	1.77	0.20
Valine, leucine and isoleucine degradation[1]	27	0.62	1.75	0.23
Krebs TCA cycle[1]	22	0.63	1.74	0.25
Proteasome pathway[2]	17	0.77	1.73	0.25

* Biological processes based on [1]GenMapp, [2]Biocart and [3]SigmaAldrich.

† q-value>0.25 was used as cut-off as advised by GSEA.

N, Number of genes in the geneset.

ES, enrichment score for the gene set.

NES, normalized enrichment score.

q-values, FDR derived q-values.

Supplemental table 6.5 Complete dataset of FOS affected colonic genes is available at www.foodbioactives.nl.

Dankwoord

Dankwoord

Alle mensen die een bijdrage hebben geleverd aan het totstandkomen van mijn proefschrift: Heel erg bedankt. Een aantal mensen wil ik speciaal bedanken.

Jaap, Ingeborg, Roelof en Martijn, bedankt voor de goede samenwerking binnen dit project. Vier "leermeesters" met ieder een eigen kijk op de zaak leerde mij dat een onderzoek van veel kanten benaderd kan worden, maar vooral leerde het me een eigen kijk op de zaak te vormen. Jaap, ondanks je drukke agenda stond je deur altijd open en wist je m'n stukken in rap tempo te corrigeren. Geweldig hoe jij uit de wirwar van genexpressie data altijd weer een duidelijk schemaatje weet te toveren. Naast je gedrevenheid voor het onderzoek bewonder ik je aandacht voor de onderzoeker. Door jouw inzicht in persoonlijke zaken, en het besef dat ook dat belangrijk is voor succesvol onderzoek, heb ik de afgelopen vier jaar veel over mijzelf geleerd. Ingeborg, we hadden steeds vaker overleg op afstand. Gelukkig is de afstand Ede-Wageningen klein, en zo voelde het ook. Door je scherpe focus zag jij de gaten in m'n redeneringen en gaf me altijd goede suggesties waar ik weer mee vooruit kon. Jouw grondige beoordeling van manuscripten leverde mooie track change-manuscripten op, ik zal je "mooie zinnen" erg gaan missen!
Roelof, officieel sta je niet in het lijstje begeleiders, maar de praktijk was anders. Jouw kritische vragen leverde interessante en leerzame discussies op. Jouw grenzeloze kennis over de darm waren belangrijk voor de vertaalslag van genenlijsten naar relevante processen in de darm. Martijn, van een afstand bewaakte je de voortgang van het project. Je scherpzinnige opmerkingen en vragen zetten mij altijd goed aan het denken. Je stuurde het project, en mij, de goede kant op, bedankt!

Evelien, jouw bijdrage aan dit project was onmisbaar! Je zou "stand in" kunnen zijn tijdens m'n promotie. We klikken erg goed op aanpak van proeven maar ook op persoonlijk vlak kan ik met jou altijd even bijkletsen. Zullen we maar gewoon de rest van onze carrière een duobaan nemen? Carolien, jouw nuchterheid tijdens het stressvolle spotten van de arrays, drukke dierproeven of moeizame schrijfmomenten was heel welkom, de expressie van m'n stress-genen schoot omlaag! Annelies, je zat officieel niet in m'n project, maar door je grote betrokkenheid leek dat wel zo. Je hebt waarschijnlijk ook als één van de weinigen m'n proefschrift zo grondig gelezen! Bedankt voor alle hulp en humor op het lab! Susan, door jou werd hoofdstuk 3 een feit. Naast je wetenschappelijke bijdrage was het vooral ook erg gezellig om met jou samen te werken!

Ik heb met veel plezier op het RIKILT gewerkt. Vooral de tijd op kamer 1.53 was super; die kamer blijft toch het centrum van gezelligheid! Ik zou met de belevenissen en verhalen van de afgelopen vier jaar met gemak een boek kunnen vullen (sorry één boek dit jaar is toch even genoeg)! Naast de vele tijdelijke gasten was er altijd een harde "blij dat je blijft" kern. Evelien en Annelies, ik had de zakdoekjes bijna echt nodig toen ik met m'n kar naar boven vertrok. Marjolein, Susan en Melissa bedankt voor de ontelbaar vele gezellige koffie-momenten, en voor een luisterend oor na succesvolle resultaten, maar vooral ook bij dipjes. Vincent, Evert, Nicole, Ping, Jeroen, Hakan en Marjoke voor iedere data analyse vraag wist ik jullie te vinden. Hakan, ook al kende ik alle schroefjes van de scanner steeds beter, scannen zonder jou in de buurt durfde ik bijna niet meer aan, dank je voor je vele reddingen! Alle Food Bioactives collega's en andere Moleculaire Biologie collega's van het RIKILT, bedankt voor alle discussies, hulp en adviezen, maar vooral ook de goede sfeer, gezellige koffie pauzes en niet te vergeten de succesvolle deelnames aan de WE-day zeskamp. Jan en Ruud bedankt voor alle computer ondersteuning.

De collega's van de Health & Safety groep van het NIZO wil ik bedanken voor alle actieve literatuur en werkbesprekingen. Deze besprekingen waren heel belangrijk voor m'n "gut- feeling" over voeding en de darm. Sandra en Mischa jullie gaven het stokje met veel hulp aan mij door, bedankt! Nizo-Aio's Gabriele, Marleen en Marloes, wanneer gaan we een volgend restaurant in Utrecht uittesten? Denise, Hans, Corinne, Carolien, Arjan, Johan en Fanny bedankt voor hulp, gezellige koffie pauzes en discussies.

Wilma, Judith, Pam en andere CKP-ers bedankt voor de verzorging van de vele ratten wat tijdens een Salmonella-FOS proef niet altijd een pretje was. Door jullie liepen de proeven en de secties altijd gesmeerd.

Collega's van TI Food and Nutrition programma 1, bedankt voor de kennismaking met het onderzoek in de humane voeding tijdens de donderdag ochtend bijeenkomsten. Het was leerzaam en leuk om in zo'n multidisciplinaire omgeving m'n onderzoek te presenteren en te bediscussiëren. Rianne, Hannie en andere TIFN-medewerkers bedankt voor alle hulp.
De AiO's van Humane voeding leerde ik in Australië goed kennen, deze geweldige ervaring was een goede basis om niet als vreemde op de vakgroep rond te lopen. Jan, ook al kwam de onderwijs begeleiding me tijd-technisch nooit goed uit, de practicum middagen vond ik een leuke onderbreking van m'n onderzoek. Geert, bedankt voor de goede samenwerking op de vele werkplekken die we hadden! Of het nu RIKILT, RIVM, Universiteit Maastricht of de WUR was, we vonden altijd wel een computer om aan te werken! Heleen, onze gezamenlijke eindsprint en de promotie-hotline gaven me een rustig gevoel; we zouden toch zeker niet beide iets heel belangrijks vergeten?

Naast alle collega's wil ik ook vrienden en familie bedanken. Zonder de broodnodige sportieve en sociale uitspattingen zat ik hier niet zo ontspannen dit dankwoord te typen. Bedankt voor al jullie interesse en aanmoediging, het evenwicht was even kwijt de afgelopen maanden, maar dat ga ik zo snel mogelijk weer herstellen.
Ineke, na een heerlijk weekendje Herkenbosch kon ik er altijd weer even tegenaan!
Pap bedankt voor je enthousiasme en interesse in m'n werk, en mam toch een beetje in je voetsporen getreden! Vanaf nu geen zorgen meer over promotie-stress, ook voor jullie een opluchting!
Joost, deze ereplaats is voor jou, de eindsprint was niet mogelijk geweest zonder zo'n geweldige thuisbasis, bedankt voor je onvoorwaardelijke steun de afgelopen jaren.

Bedankt!
Wendy

About the author

CV

Wendy (Gwendolyn Cornelia Hendrika) Rodenburg was born on July 23th 1975 in Bodegraven, the Netherlands. In 1994 she started the study Biology at Utrecht University, with specialization Fundamental Biomedical Sciences. As part of this study she conducted three trainee-projects. In the Immunotoxicology group at RITOX, Utrecht University, she examined the immune response in mice upon exposure of skin to irritants and allergens. At the Bioscience division, Unilever Research, Colworth UK she studied the effect of ageing and sun exposure on Matrix metalloproteinase I mRNA level, protein level and enzyme activity in human skin fibroblasts. In the Department of Molecular Cell Biology, Faculty of Biology, Utrecht University she examined potential blood biomarkers for obesity in serum of morbid obese human subjects. After receiving her MSc degree in 1999 she joined the Department of Molecular Cell Biology where she studied the regulation of Protein Kinase B by Protein Kinase C ζ in insulin signalling. In 2002 she was appointed by the division of Human Nutrition of Wageningen University to conduct the PhD-project described in this thesis. This project was part of TI Food and Nutrition and was carried out at the Food Bioactives Group of RIKILT- Institute of Food Safety and at the Health and Safety Department of NIZO Food Research under supervision of Dr J. Keijer, Dr. ir. I.M.J. Bovee-Oudenhoven and Prof. dr. M.B. Katan.

Publications

Rodenburg W., Bovee-Oudenhoven, I.M.J., Kramer E., van der Meer R., Keijer J.
Gene expression response of the rat small intestine following oral Salmonella infection
Physiological Genomics, 30(2), 123-33, 2007

Rodenburg W., Keijer J., Kramer E., Roosing S., Vink C., Katan M.B., van der Meer R., Bovee-Oudenhoven, I.M.J.
Salmonella induces prominent gene expression in the rat colon
BMC Microbiology, 7(1):84, 2007

Rodenburg W., Heidema A.G., Boer J.M.A., Bovee-Oudenhoven, I.M.J., Feskens E.J.M., Mariman E.C.M., Keijer J.
A framework to identify physiological responses in microarray based gene expression studies: selection and interpretation of biologically relevant genes
Submitted

Rodenburg W., Keijer J., Kramer E., Vink C., van der Meer R., Bovee-Oudenhoven, I.M.J.
Impaired intestinal barrier function by dietary fructo-oligosaccharides (FOS) in rats is associated with increased mitochondrial gene expression
Submitted

van Ampting M., **Rodenburg W**., Vink C., Kramer E., Schonewille A.J., Keijer J., van der Meer R., Bovee-Oudenhoven, I.M.J.
Ileal mucosal and faecal pancreatitis associated protein is increased during Salmonella infection in rats and is associated with infection severity
Submitted

Training and supervision plan

Courses

Ecophysiology of the GI-tract	VLAG, Wageningen	2003
Masterclass Nutrigenomics	VLAG, Wageningen	2003
Bioinformatics (BIT I)	WUR, Wageningen	2003
Statistics for microarray experiments	UU, Utrecht	2003, 2005
Proefdierkunde	UU, Utrecht	2003
English Scientific Writing	CENTA, Wageningen	2005
Debating skills for PhD-students	TIFN, Wageningen	2006

Meetings

NWO-nutrition	NWO, Papendal	2004 - 2006
Darmendag	IOP, NUTRIM	2004 - 2007
Functional genomics of Host-Pathogen interactions	CSHL, Hinxton, UK	2004
NUGO week	NUGO	2004, 2005
NUGO WP Gut Health meeting	NUGO, Sienna, Italy	2005
Gasto-intestinal tract	FASEB, Snowmass Village, USA	2005
International conference on Nutrition and Intestinal Health	Kaiserslautern, Germany	2007

General

PhD study tour Australia	WUR, Human Nutrition	2003
RIKILT Molecular Biology Research presentations	RIKILT	2003-2007
Journal club NIZO Food Research	NIZO	2003-2007
TIFN Seminars	TI Food and Nutrition	2003-2007