UNIVERSITY OF TRENTO - Italy

International PhD Program in Biomolecular Sciences

Centre for Integrative Biology

XXVII Cycle

COSBI

Modulation of Murine Intestinal Gene Expression by Dietary Fat: Spatial Variation and Saturation Effect

Tutor

Dr. Marco Scotti

GEOMAR Helmholtz Centre for Ocean Research Kiel

Germany

Advisor

Dr. Melissa J. Morine

The Microsoft Research – University of Trento Centre for Computational Systems Biology

Ph.D. Thesis of

Tenzin Nyima

The Microsoft Research – University of Trento Centre for Computational Systems Biology

Academic Year 2014-2015

To My Parents

Declaration

I hereby confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Tenzin Nyima

Date: 31 January 2016

Place: Rovereto

Acknowledgements

I had a remarkable research experience during my PhD journey. I am deeply indebted to my Tutor Dr. Marco Scotti and thesis Advisor Dr. Melissa J. Morine for their invaluable guidance, and for giving me this wonderful opportunity to work on this project. This PhD would not have been possible without their expertise, dedication and enthusiasm. Also, I am very thankful to Dr. Marco Scotti for his patience, support and for being an excellent supervisor. This thesis has been thoroughly reviewed and proofread by him, for which I am very grateful.

I am thankful to our collaborators Dr. Guido J. E. J. Hooiveld from Wageningen University (The Netherlands) and Prof. Michael Müller from University of East Anglia in Norwich (UK) for allowing us to analyze their data. And, I would like to thank them for hosting me at their lab (Division of Human Nutrition, Wageningen University) during my period abroad. I had a gratifying time working under their guidance, which helped me gain new perspective on my research.

Research on this thesis was made possible by the fellowship provided by COSBI, for which I am truly grateful. Also, I would like to extend my thanks to COSBI for providing me the resources necessary to complete this project. I had an enjoyable time working with my brilliant colleagues at COSBI and I would like to thank them for their support.

I am very thankful to our PhD school for providing the academic support. I would like to thank Prof. Paolo Macchi and Betty Balduin for their patience and advice throughout the academic years.

I am grateful to Dr. Thanh-Phuong Nguyen and Dr. Tamás Korcsmáros for kindly agreeing to referee my thesis.

Finally, I would like to convey my heartfelt thanks to my parents and family members and friends for supporting my education throughout my life. Nothing would be possible without their consistent support and faith in me.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ABC ATP-binding cassette

ACAT acyl-CoA cholesterol acyltransferase

ADA Adenosine deaminase

APR Acute-phase response

Birc5 Baculoviral IAP repeat-containing

BMI Body mass index

CCK Cholecystokinin

ChEt Cholesterol esterase

DEG Differentially expressed genes

FAT Fatty acid transport protein

FATP4 Fatty acid transport protein

FXR Farnesoid X receptor

GO Gene Ontology

GOBP Gene Ontology Biological Processes

GPCR G protein-coupled receptors

GPX Glutathione peroxidase

HD5 Alpha defensin

HF High-fat

IAP Intestinal alkaline phosphatase

I-FABP Intestinal fatty acid binding protein

KEGG Kyoto Encyclopedia of Genes and Genomes

LDLs Low-density lipoproteins

LF Low-fat

L-FABP liver fatty acid binding protein

LPL Lipoprotein lipase

LPS Lipopolysaccharide

MM Mismatch

MTP Microsomal triglyceride transport protein

NPC1L1 Niemann-Pick C1 like 1 protein

NUSE Normalized Unscaled Standard Error

PLA2 Phospholipase A2

PLM Probe Level Model

PM perfect match

Reg3A Regenerating islet-derived 3 alpha

RLE Relative Log Expression

RMA Robust multi-array average

Serpina1b Serine protease inhibitor

SR-B1 class B Scavenger receptor

SREBP Sterol regulatory element binding protein

TLR4 Toll-like receptor

TNF Tumor necrosis factor

UPC Universal exPression Code

VLDL Very low-density lipoprotein

ZO Zonula occludens

Chapter 1

Introduction

1.1 Effects of dietary fat in the small intestine

Overconsumption of high caloric diet and sedentary lifestyle promote overweight and obesity (Balwierz et al. 2009; Cheung & Mao 2012; Desmarchelier et al. 2012; Walley et al. 2006; Zhang et al. 2009). From 1980 to 2013, the percentage of adults with body mass index (BMI) of 25 kg/m^2 or greater increased from 28.8% and 29.8% to 36.9% and 38.0% in men and women, respectively (Ng et al. 2014). Obesity is associated with multiple comorbidities such as insulin resistance and type 2 diabetes (Khaodhiar et al. 1999). The incidence of both obesity and overweight were estimated to cause 3.4 million deaths and 3.8% of disability-adjusted life-years (i.e., the number of years lost due to premature death, poor health or disability) globally in 2010 (Ng et al. 2014; Lim et al. 2012). The cost for obesity worldwide amounts to $2 trillion annually (Dobbs et al. 2014). Diet and lifestyle are relevant for determining the health status of individuals, but also gut microbiota plays a key role in health and disease of human (O'Hara & Shanahan 2006). For example, it modulates lipid metabolism by modifying bile acid metabolism (Martin et al. 2007) and influences energy balance by promoting energy harvest from the diet and energy storage into the host (Bäckhed et al. 2004, 2007). Diet plays an essential role in shaping intestinal microbiota and high-fat diet exerts a major influence on it (Hotamisligil 2006; de La Serre

et al. 2010). An obese individual is characterized by lower bacterial diversity with increased proportion of harmful (Firmicutes) in contrast with health-promoting (Bacteroidetes) gut microbiota; this condition is termed as dysbiosis (Bäckhed et al. 2004; Clemente et al. 2012; Fleissner et al. 2010; de La Serre et al. 2010; Turnbaugh et al. 2008, 2009). Such reduced bacterial diversity and the presence of dysbiotic conditions were also observed in individuals consuming western style diet (De Filippo et al. 2010; Yatsunenko et al. 2012). A typical western diet consists of approximately 40% of calories from fat although the recommended range is 20-30% (Institute of Medicine 2005; Niot et al. 2009; Werner et al. 2000). Individuals that experience weight loss or follow non-western style diets (i.e., rich in fiber content) exhibit an opposite pattern of bacterial diversity if compared to obese individuals (Bäckhed et al. 2004; De Filippo et al. 2010; Ley et al. 2006; Turnbaugh et al. 2008; Yatsunenko et al. 2012). Obesity and related metabolic complications cause chronic or low-grade inflammation (Hotamisligil 2006). Chronic inflammation is triggered by the increased proportion of lipopolysaccharide containing gut microbiota, also known as metabolic endotoxemia (Cani et al. 2007). Increased endotoxemia is observed during excessive high-fat intake, which dysregulates inflammatory signal and leads to weight gain and diabetes (Cani et al. 2007). Dysregulated inflammatory signal is responsible for the high-fat mediated hyperphagia and obesity (Hotamisligil 2006; de La Serre et al. 2010).

The small intestine functions as a mediator between the internal and the external environment of the body, and is essential for the digestion and the absorption of fat (Kondo et al. 2006; Petit et al. 2007). Mice fed with high-fat diet exhibit prominent up-regulation of lipid metabolism related genes in the small intestine and show significant predisposition to obesity (Kondo et al. 2006; de Wit et al. 2011). The intestine displays regionalized distribution in functionality and its longitudinal axis is divided into three sections (from the most proximal to the most distal section): duodenum, jejunum, and ileum. Jejunum

consists of maximal absorptive surface and the size of lumen decreases from the duodenum to the ileum (DeSesso & Jacobson 2001). When mice are fed with dietary fat gradient (i.e., 10%, 20%, 30%, or 45% kcal from fat), the small intestine displays a buffer capacity in fat handling up to an intake of fat that represents 20% of energy demand (de Wit et al. 2011). Beyond such a 20% threshold, a shift towards an obese phenotype is observed (i.e., significant increase of body weight). The most striking effect in fat-induced differential gene expression is detected when the relative contribution of fat (compared to total energy demand) is between 30% and 45% (de Wit et al. 2011).The small intestine displays adaptability to the lipid content by enhancing cell proliferation and up-regulating the genes involved in its uptake (Niot et al. 2009; Petit et al. 2007). For example, the up-regulation of fatty acid transport protein (FAT)-4 promotes the uptake of long-chain fatty-acids across intestinal enterocytes (Herrmann et al. 2001). Permeability across intestine occurs either through the enterocytes (i.e., transcellular) or via the intraepithelial tight junctions (i.e., paracellular; Balzan et al. 2007). Enhanced paracellular permeability is observed during excessive high-fat intake, obesity and diabetes (Balzan et al. 2007; Cani et al. 2008; de La Serre et al. 2010). This is due to the modulation of gut microbiota during high-fat overconsumption. Such modulation modifies bile acid metabolism, thus leading to the breach in intestinal permeability by up-regulation of farnesoid X receptor (FXR) and tumor necrosis factor (TNF) while repressing the expression of tight junction proteins (zonula occludens-ZO-1, occludin, claudin-1, and claudin-3; see Cani et al. 2008; Martin et al. 2007; Stenman et al. 2012; Suzuki & Hara 2010). Disruption in intestinal permeability consequently increases metabolic endotoxemia that triggers inflammation and metabolic disorders. Therefore, investigation of the gene expression patterns in the small intestine in response to dietary fat intake is of key interest for better understanding the relationships between fat mediated modification of intestinal lipid metabolism and susceptibility to obesity (Kondo et al. 2006; Petit et al. 2007; de Wit et al. 2011). Figure 1.1 illustrates the

complex interplay that characterizes dietary fat overconsumption, gene expression, microbiota, and disease onset in the small intestine.

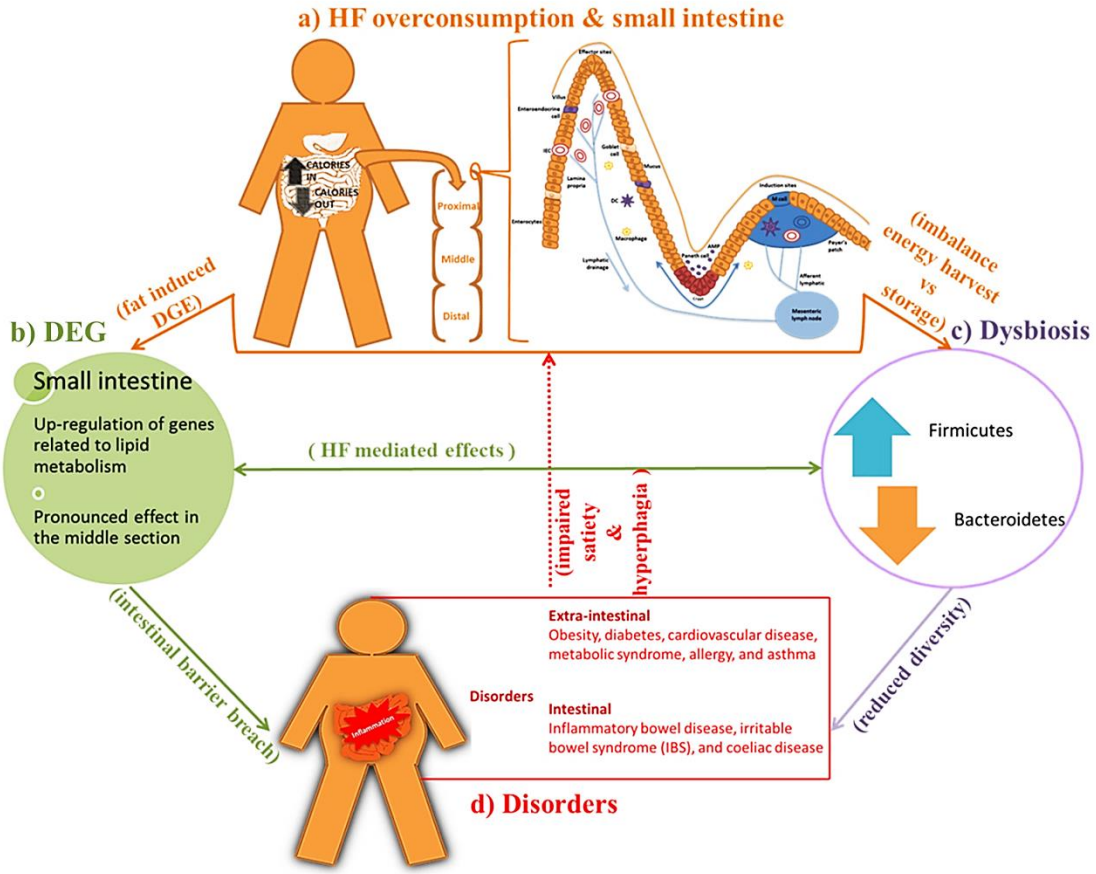

Figure 1.1 Consequences of high-fat overconsumption on small intestine. (a) Excessive consumption of high-fat diet with minimal physical activity represents a major contributing factor to the global obesity epidemic. High-fat diet exhibits section-specific impacts on the small intestine. The intestinal epithelial cells consist of absorptive enterocytes, goblet cells, Paneth cells, enteroendocrine cells and microfold cell. The maximum absorption and digestion occurs in the small intestine due to the presence of folds (plicae), depressions (crypts) and fingers like projections (villi). The magnified diagram of the structure of intestinal barrier in (a) is adapted from (Garrett et al. 2010; Mowat & Agace 2014; Mowat 2003; Peterson & Artis 2014). (b) In the small intestine of mice, a high-fat diet overexpresses genes involved in the lipid metabolism (de Wit et al. 2011). However, these genes exhibit limited change in other metabolic organs such as the liver, muscle, and white adipose tissue (Kondo et al. 2006). (c) High-fat overconsumption also impacts the microbiota; this promotes the dysbiotic condition during which Firmicutes increase at the cost of Bacteroidetes (Bäckhed et al. 2004; Turnbaugh et al. 2008, 2009). Such condition reportedly increases energy harvest from the diet by the microbiota and impairs the balance between energy harvest and storage (Bäckhed

et al. 2004, 2007). Both modulation of fat metabolism (b) and disruption of microbiota homeostasis (c) are characterized by intestinal barrier breach and reduced bacterial diversity (whilst supporting the dysbiotic condition). (d) Such condition (i.e., dietary fat mediated breach in intestinal barrier and dysbiosis) triggers inflammatory cascade, which promotes extra-intestinal and intestinal disorders (Carding et al. 2015; Fasano & Shea-Donohue 2005).

1.2 Small intestine morphology and physiology

The small intestine displays spatial differences in morphology and physiology. Its absorptive surface is greatly maximized by the presence of folds (plicae), depressions (crypts) and fingers like projections (villi; DeSesso & Jacobson 2001). Intestinal epithelial cells are lined by a single layer of columnar epithelium (they mainly refer to absorptive enterocytes, but also include goblet cells, Paneth cells, enteroendocrine cells and microfold cells) while intraepithelial cells are sealed together by tight junctions (Figure 1.1a). The intestinal epithelium renews every 4 days in human and 3 days in mice (Kaminsky & Zhang 2003; Petit et al. 2007). Among the three sections of the intestinal epithelium (i.e., duodenum, jejunum and ileum), jejunum displays the maximal surface area; the height of villi decreases from jejunum to the ileum, and lumen size decreases from proximal duodenum to distal ileum (DeSesso & Jacobson 2001). The small intestine displays high variability in gene-expression and this is due to three factors: (1) cell-type differences; (2) changes along the longitudinal axis; (3) differences in developmental stages. George et al. (2008) found high expression levels of innate antimicrobials (alpha defensin 5 - *Hd5*, and regenerating islet-derived 3 alpha - *Reg3A*) exclusively in the ileum-crypt cells. The lipid sensing mechanism of the upper intestine suppresses glucose production to maintain energy balance and regulate food intake (Wang et al. 2008). Adenosine deaminase (ADA) is activated significantly in the duodenal epithelium of mice during the suckling-weaning transition (2-3 weeks; Dusing et al. 2000).

Among the macronutrients, gastric emptying is faster for carbohydrates and slower for fats. The transit time for chyme (a partially digested food from the stomach) in the small intestine takes about 3-4 hours. An increase in transit time enhances absorption, especially for poorly absorbed substances (except anticholinergic; DeSesso & Jacobson 2001). The intestinal morphology shows adaptability to the nutritional status. Along with the decline in cell proliferation, fasting leads to gradual atrophy of rat intestinal epithelium (Dunel-Erb et al. 2001). Refeeding activates cellular proliferation and dietary lipids are the strongest stimulators (Buts et al. 1990). In mice, dietary fat-content mediated intestinal adaptation occurs through two complementary events: (1) fatty-diet enhances the height of the villi and shows strongest effect on the intestinal trophism (i.e., crypt-to-villus migration, during epithelial cell renewal); (2) high-fat diet leads to up-regulation of dietary-fat responsive genes that play a significant role in the intestinal fatty acid uptake (i.e., fatty acid transport protein 4 - FATP4, fatty acid transport protein - FAT/CD36, intestinal fatty acid binding protein - I-FABP, and liver fatty acid binding protein - L-FABP) and lipoprotein secretion (i.e., microsomal triglyceride transport protein - MTP, and apolipoprotein - apoA-IV; (Petit et al. 2007). For example, FATP4 up-regulation promotes long-chain fatty acid uptake by the enterocytes (Herrmann et al. 2001).

The activity of the small intestine is regulated via a complex interplay of different organs. The gut-brain-liver axis, together with the oral cavity, provides cumulative sensory inputs to maintain metabolic health. Both rodents and humans have developed intestinal lipid-sensing mechanisms (through vagal afferent neurons in the intestinal walls) that regulate food intake and maintain energy homeostasis (Ding & Lund 2011; de Lartigue et al. 2011). In rodents, a raise in lipid concentration in the upper part of small intestine lowers glucose production through the intestine-brain-liver neurocircuitry (Wang et al. 2008). Such signal mediates satiety via the cholecystokinin (CCK) pathway and its impairment promotes hyperphagia (Ding & Lund 2011). High-fat fed rats (i.e., with lard-oil-enriched diet)

increase caloric intake rapidly and, after three days, acquire defects in brain's lipid sensing mechanism (defects in glucose production; Pocai et al. 2006; Wang et al. 2008). Wang et al. (2008) have proposed that the lipid-sensing mechanism of the intestine provides one of the first lines of metabolic defense against nutrient excess by inhibiting food intake and maintaining energy balance. Furthermore, reduced satiety during high-fat feeding can be due to: (1) oral sensory input (this is because rats on high-fat diet reduce oral fat sensitivity with decreased expression of oral CD36 transporter; Newman et al. 2013); (2) malfunction in gut nutrient sensors like G protein-coupled receptors (GPCR, expressed in enteroendocrine cells; Janssen & Depoortere 2013)(Janssen & Depoortere 2013). On a high-fat diet, lipid metabolism related genes are prominently over-expressed in the small intestine while a limited change can be detected in the liver, muscle, and white adipose tissues (Kondo et al. 2006; see Figure 1.1b).

1.3 Lipid absorption and digestion in small intestine

Dietary fat belongs to a variety of lipid classes and each class exhibits unique metabolic phases in different cellular environments. The lipid classes are categorized based on the chain length or the interaction with water. Based on the chain length (C2 to C24), fatty acids have been grouped into four classes: saturated, monounsaturated, n-6 polyunsaturated and n-3 polyunsaturated (Nassir & Abumrad 2009). The solubility of lipids in water plays a key role in their uptake across membranes and, based on interaction with water, lipids are categorized as non-polar (e.g., cholesteryl oleate) and polar (e.g., triacylglycerols; Phan & Tso 2001; Schulthess et al. 1994; Werner et al. 2000). In a western diet, almost 40% of the calories are contributed by the dietary fat (i.e., about 90-95% triacylglycerols , 4.5% phospholipids and sterols; Niot et al. 2009; Werner et al. 2000). Most of the phospholipids in intestinal lumen are from the bile, while only a small amount comes from the diet (Phan & Tso 2001; Werner et al. 2000). Most of the dietary sterol is cholesterol (animal origin)

and only a small amount from plants (mainly beta-sitosterol). Beta-sitosterol is efficiently released from the enterocytes into the intestinal lumen, due to intestinal half-transporters ATP-binding cassette (ABC) transporters G5 (ABCG5) and G8 (ABCG8; see Berge et al. 2000). Mammalian cells cannot synthesize a specific group of lipids called essential fatty acids, which must be obtained from the diet. The fat-soluble vitamins (A, D, E and K) are essential for the normal functioning of cells and organs. The absorption rates are higher for vitamins A, D and K compared to vitamin E. It has been suggested that competition between these fat-soluble vitamins can exist for what concerns intestinal absorption and transport (Phan & Tso 2001; Werner et al. 2000). The intestinal lipid absorption occurs via two successive events: intraluminal and intracellular processes (Werner et al. 2000).

1.3.1 Intraluminal lipid absorption

Multiple physico-chemical changes involve dietary lipids before the uptake from the intestinal lumen to the enterocytes (Werner et al. 2000). Intraluminal lipid absorption includes: (1) emulsification and lipolysis; (2) solubilization; (3) uptake of lipolytic products by enterocytes (Phan & Tso 2001; Werner et al. 2000).

(1) Emulsification and lipolysis. Dietary fat digestion begins in the stomach with mechanical emulsification and partial triacylglycerols hydrolysis by the acid lipases (lingual and gastric lipases; Phan & Tso 2001; Werner et al. 2000). Most of the triacylglycerols digestion occurs in the duodenum and the first part of the jejunum by the combined action of bile and pancreatic lipase (Borgström 1977; Werner et al. 2000). Pancreatic lipase hydrolyzes triacylglycerols into monoglycerides and fatty acids (Nassir et al. 2007). When triacylglycerols droplets are covered with bile salts, pancreatic lipase requires the cofactor pancreatic co-lipase for adequate triacylglycerols hydrolysis (Werner et al. 2000). The digestion of phospholipids occurs mainly in the duodenum by the pancreatic phospholipase A2 (PLA2) which yields fatty acids and lyso-

phosphatidylcholine (Iqbal & Hussain 2009; Werner et al. 2000). Dietary cholesterol is mostly available in the form of free sterol, with only 10-15% cholesterol ester. Before absorption, cholesterol esters must be hydrolyzed in the duodenum by pancreatic cholesterol esterase (ChEt; see Iqbal & Hussain 2009; Phan & Tso 2001; Werner et al. 2000). Pancreatic ChEt in humans shows wide specificity (it acts on triacylglycerols, cholesterol esters and phosphoglycerides) and its activity is considerably amplified by the presence of bile salts (Lombardo et al. 1980; Phan & Tso 2001; Werner et al. 2000).

(2) Solubilization. The luminal content (nutrients and bacteria) of small intestine is separated from the intestinal epithelial cells by the unstirred water layer, which is secreted by the mucin producing cells (e.g., goblet cells; Turner 2009). In order to pass through the unstirred layer, dietary fats need to be solubilized into mixed micelles by the bile acids (Wang et al. 2013). Mixed micelles act as transporters by delivering emulsified dietary fat to the apical enterocytes for the uptake (Iqbal & Hussain 2009; Schwenk et al. 2010; Wang et al. 2013).

(3) Uptake of lipolytic products by enterocytes. The diffusion of fatty acids across the enterocytes was believed to depend on simple passive-diffusion (Iqbal & Hussain 2009; Schwenk et al. 2010; Wang et al. 2013). However, fatty acid uptake occurs mainly by membrane-associated FABP (Schulthess et al. 1994; Wang et al. 2013). There are three main mechanisms for the uptake and transport of fatty acids across the enterocytes (Wang et al. 2013). These mechanisms govern fatty-acid uptake and are mediated by membrane-associated FABP; they refer to: (1) CD36 (also known as fatty acid translocase); (2) FATP4; (3) combined action of Niemann-Pick C1 like 1 protein (NPC1L1), ABCG5 and ABCG8 (Wang et al. 2013). CD36 is a multifunctional protein homologous to the class B scavenger receptor (SR-B1; Niot et al. 2009). It is abundantly expressed at the major site of fat absorption, the duodeno-jejunum (Poirier et al. 1996). Rat and human studies show a precise localization of CD36 in the apical enterocytes. The magnitude of CD36 expression

exhibits a clear positive relationship with dietary-lipid content, with significant up-regulation in rats fed with high-fat diet (Iqbal & Hussain 2009).CD36 is also up-regulated in case of genetic obesity and diabetes mellitus (Greenwalt et al. 1995; Iqbal & Hussain 2009). CD36-null mice show abnormal lipid processing with the accumulation of lipids in the proximal small intestine; thus, CD36 plays a vital role in both secretion and clearance of intestinal lipoproteins (Drover et al. 2005). The activity of CD36 is enhanced by the plasma membrane-associated FABP (FABPpm) that facilitates medium- and long-chain fatty acid uptake across the apical membrane of the enterocytes (Wang et al. 2013). After uptake, fatty acids are bound by the cytoplasmic FABP (FABPc) and enter various metabolic pathways (e.g., oxidation in mitochondria; Georgiadi & Kersten 2012; Wang et al. 2013). FATP4 transports medium- and long-chain fatty acids (Wang et al. 2013). FATP4 is more expressed in mature apical enterocytes than in crypts cells (where its expression is low or almost negligible); higher expression levels are usually found in jejunum than duodenum (Stahl et al. 1999). Its relevance in the *in vivo* fat absorption remains unclear but a role in the triacylglycerols processing is suggested. This is because FATP4-null mice show increased accumulation of triacylglycerols in the enterocytes (Shim et al. 2009). Intestinal cholesterol can also be maintained by the combined effects of NPC1L1 and ABC half-transporters (ABCG5 and ABCG8; Iqbal & Hussain 2009; Wang et al. 2013). NPC1L1 facilitates sterol influx at the apical enterocytes, whereas cholesterol and plant sterol efflux into the intestinal lumen is actively promoted by ABC half-transporters (Wang et al. 2013). NPC1L1 deficient mice show reduced cholesterol absorption (Altmann et al. 2004). Moreover, the deletion of *Cd36* in mice results in up-regulation of NPC1L1 in the middle and distal sections of small intestine (Nassir et al. 2007). Cholesterol absorption is minimally affected in *Abcg5* and *Abcg8* deficient mice that show reduced biliary cholesterol secretion and increased absorption of dietary plant sterol (Yu et al. 2002).

1.3.2 Intracellular lipid absorption

After crossing the apical enterocytes, dietary lipids are directed into the endoplasmic reticulum where they are converted into the bio-inert triacylglycerols. Conversion into triacylglycerols is needed since lipolytic products (especially free fatty acids present in high concentrations) are toxic to the intestinal cells that are thus exposed to be damaged (Mansbach & Nevin 1998). As a countermeasure to toxic damages, FABP aids in the transport of absorbed fatty acids (FABP is more available in the villi than in the crypts and more expressed in the jejunum than in the ileum; Ockner et al. 1972; Ockner & Manning 1974). In the proximal section, FABP is expressed in two forms: L-FABP and I-FABP. The fate of unbound fatty acids from the intestinal cytosol mainly follows two ways: (1) L-FABP activates a storage mechanism for the free fatty acids and monoacylglycerol; (2) when L-FABP is overwhelmed by excess free fatty acids fluxing into the enterocyte, then free fatty acids and monoacylglycerol are rapidly esterified into triacylglycerols (Mansbach & Dowell 2000). However, Vassileva et al. demonstrated that I-FABP is not essential for the dietary fat absorption but rather maintains energy homeostasis; I-FABP performs such function by acting as a lipid-sensing component that alters body weight gain in a gender-specific manner (Vassileva et al. 2000).

The absorbed cholesterol is dispatched into the free cholesterol pool inside the enterocytes and is mostly transported into the lymphatic system as cholesterol ester, which is esterified by the pancreatic ChEt and acyl-CoA cholesterol acyltransferase (ACAT; Clark & Tercyak 1984). Higher activity of ACAT has been observed in the jejunum compared to the ileum (Phan & Tso 2001) and two types of ACAT have been identified: ACAT-1 and ACAT-2. The expression of ACAT-2 is more restricted to small intestine and liver. Its role in the intestinal cholesterol absorption and esterification is demonstrated, with ACAT-2 knockout mice that are resistant to diet-induced hypercholesterolemia (Buhman et al. 2000). In the smooth endoplasmic reticulum, newly synthesized triacylglycerols and cholesterol esters

are packaged into a lipoprotein particle called chylomicron (Werner et al. 2000). These nascent chylomicrons are covered with phospholipids, cholesterol and apolipoproteins (apoA-I, apoA-IV and apoB-48). The gradual rise of fat absorption and triacylglycerols resynthesis increases the size and number of lipoprotein particles, which are packaged in vesicles filled with pre-chylomicrons. After modification of the pre-chylomicrons into mature chylomicrons in the Golgi apparatus, these vesicles translocate to the lateral enterocytes to be exocytosed into the interstitium and the lymph. Then, they enter the thoracic duct via the mesenteric lymph duct and finally reach the systemic circulation (Werner et al. 2000). During postprandial period, an increase in circulating chylomicrons induces the peak of triglycerides. This peak in plasma triacylglycerol is highest in obese and diabetic individuals (Georgiadi & Kersten 2012). Circulating chylomicrons are further hydrolyzed by the endothelial lipoprotein lipase (LPL) and then fatty acids are distributed to the peripheral tissues (muscles and adipose tissues; Georgiadi & Kersten 2012; Iqbal & Hussain 2009; Niot et al. 2009).

1.4 High-fat diet and microbiome in intestine

In the human body, most of the interactions between host and microbial cells occur along the mucosal surface, and gut is densely populated by commensal and symbiotic bacteria (O'Hara & Shanahan 2006; Shen et al. 2013). The ensemble of microbial cells that are harbored by the human body constitutes the microbiota and all genes encoded by them are the microbiome. In the gut, the microbiome includes 3.3 million non-redundant genes (i.e., they represent an amount of genes that is 150 times larger than the human gene complement; Clemente et al. 2012; Qin et al. 2010). Metabolism and survival of these two coevolved genomes (human and microbiome) are inextricably interwoven, and microbiota performs a major role in health and disease of human (Clemente et al. 2012; O'Hara & Shanahan 2006). The gut microbiota composition is highly influenced by diet (Fleissner et

al. 2010), but its plasticity (i.e., adaptability) in response to diet is also affected by the genome of the host (Parks et al. 2013). When the balanced activity of health-promoting and deleterious/neutral gut microbiome is disrupted, a condition called dysbiosis emerges (Figure 1.1c). This promotes the development of obesity, inflammatory bowel disease, diabetes and other metabolic complications (Clemente et al. 2012; Shen et al. 2013).

The microbiota of a healthy human gut is composed of six main bacterial phyla: Firmicutes, Bacteroidetes, Proteobacteria, Actinobacteria, Fusobacteria and Verrucomicrobia (Eckburg et al. 2005; Shen et al. 2013; Zhang et al. 2009). However, more than 90% of the total gut microbiota comprises Firmicutes and Bacteroidetes (Clemente et al. 2012; Garrett et al. 2010; Shen et al. 2013; Velagapudi et al. 2010).The composition (species/strains members and their abundance) of gut microbiota is unique for each individual (Lozupone et al. 2012; Shen et al. 2013; Zoetendal et al. 1998). Bacterial diversity is lower in obese individuals, with higher abundance of Firmicutes compared to Bacteroidetes (Bäckhed et al. 2007; Clemente et al. 2012; Fleissner et al. 2010; de La Serre et al. 2010; Turnbaugh et al. 2008, 2009). In presence of weight loss there is an increased proportion of Bacteroidetes with a reduction in Firmicutes (Bäckhed et al. 2007; Ley et al. 2006; Turnbaugh et al. 2008). Dietary habits play an essential role in shaping the gut microbiota community. For example, the fecal microbiome of western populations (i.e., from US metropolitan areas) shows less bacterial diversity than the one of populations from non-western areas, and even the phylogenetic composition varies significantly according to the geographical location (when this well approximates different feeding behaviors; Yatsunenko et al. 2012). The gut microbial composition of African children (consuming fiber rich, non-western diet) is enriched in Bacteroidetes and depleted in Firmicutes if compared to European children (consuming western-style diet; De Filippo et al. 2010). In general, western style diet is associated with the reduction of bacterial diversity in the gut and promotes dysbiotic conditions.

1.5 Dietary fat intake and diseases in intestine

Obesity and other metabolic disorders such as diabetes are characterized by a chronic or low-grade inflammation (Hotamisligil 2006). A condition called metabolic endotoxemia, during which there is an increased proportion of lipopolysaccharides containing gut microbiota, is the causative factor for the associated inflammatory state (Cani et al. 2007). Excess high-fat intake increases the proportion of lipopolysaccharides, which trigger body weight gain and diabetes by dysregulating the inflammatory signal (Cani et al. 2007). Lipopolysaccharides regulate the insulin sensitivity and the onset of diabetes and obesity by binding to the CD14-dependent TLR4 complex of the innate cells (Cani et al. 2007; Wright et al. 1990). Moreover, TLR4 seems to be the molecular link among nutrition, lipids and inflammation such that its deletion partially protects mice from high-fat diet-induced insulin resistance (Shi et al. 2006).Gut microbiota has influence over both sides of the energy balance (i.e., energy harvest from the diet and energy storage into the host) and high-fat diet induced microbiota composition exhibited enhanced energy harvest from the diet (Bäckhed et al. 2004, 2007). Mice fed with high-fat diet show increased intestinal permeability with the down-regulation of the epithelial tight junction proteins ZO-1 and occludin, a condition that leads to liver inflammatory damages by increasing lipopolysaccharides levels in the portal blood (Brun et al. 2007; Cani et al. 2008). Gut microbiota is involved in this mechanism and its relevance is corroborated by the fact that when mice fed with high-fat diet are treated with antibiotic they display normal intestinal integrity and down-regulation of several inflammatory markers such as plasminogen activator inhibitor-1 (PAI-1), Interleukin 1 family (IL-1), and tumor necrosis factor alpha (TNF-α; Cani et al. 2008). The intestinal barrier disruption (a process that is triggered by environmental factors such as dietary or microbial antigens) can stimulate the immune response in extra-intestinal organs such as liver, pancreas, kidney, the skeletal system and

the brain (Fasano & Shea-Donohue 2005). Moreover, dysbiotic gut microbiota has consequences on both (1) extra-intestinal disorders (e.g., obesity, cardiovascular diseases, metabolic syndrome, asthma, and allergy) and (2) intestinal disorders (e.g., inflammatory bowel diseases like Crohn's disease, irritable bowel syndrome and coeliac disease (Carding et al. 2015; see Figure 1.1d). Many disorders display regionalized distribution along the longitudinal axis of small intestine. Microarray transcriptomic data analysis has revealed that, during dietary fat-induced development of obesity and insulin resistance, the most pronounced effects are in the middle small intestine (with most of the differentially expressed genes that are related to lipid metabolism; de Wit et al. 2008). Among the intestinal inflammatory disorders, Crohn's disease tends to be restricted to distal small intestine and colon, while coeliac disease is more characteristic of duodenum and upper small intestine (Knights et al. 2013; Meresse et al. 2012). Crohn's disease displays aberrant inflammatory response to commensal bacteria whereas coeliac disease is triggered by dietary gluten (Knights et al. 2013; Meresse et al. 2012).In patients with ileal Crohn's disease, the function of Paneth cells is compromised and results in a decrease of antimicrobial peptide production (i.e., NOD2, ATG16L1 and immunity-related GTPase family M protein 1 - IRGM-1; see Knights et al. 2013; Mowat & Agace 2014; Peterson & Artis 2014; Vaishnava et al. 2008). In patients with coeliac disease, dietary gluten enhances intestinal permeability by MyD88 dependent up-regulation of zonulin, which is followed by antigen-specific adaptive immune response (Fasano et al. 2000; Fasano & Shea-Donohue 2005; Thomas et al. 2006). The complex relationship involving different environmental factors such as diet and microbiota can alter the balance between immunity and tolerance thus playing an important role for the onset of metabolic and intestinal disorders. Dysbiotic conditions can be improved by treating the patients with antibiotic, prebiotic (increased consumption of dietary fiber) and probiotics (dietary supplements containing beneficial bacteria; see Cani et al. 2008; Everard et al. 2011; Serino et al. 2009).

Prebiotic treatment enhanced the production of GLP-1 and GLP-2 from the enteroendocrine L cells, which play a vital role in the regulation of gut barrier function and glucose homeostasis (Delzenne & Cani 2011). Oligofructose increased the abundance of *Akkermansia muciniphila*, which reversed metabolic disorders induced by high-fat diet (such as metabolic endotoxemia and insulin resistance; Everard et al. 2013). Malt counteracted the fat mediated increase in the mRNA expression of tight junction proteins and TLR in the small intestine and distal colon (Zhong et al. 2015).

1.6 Nonlinear transcriptomic response to dietary fat intake

Dietary fatty acids act as signaling molecules and influence metabolic processes by stimulation/inhibition of DNA transcription (Georgiadi & Kersten 2012). The transcriptional regulation of genes associated with the fatty acid metabolism has a pivotal role in modulating the long-term regulatory mechanism that controls lipid homeostasis. This control is executed by means of transcription factors such as sterol regulatory element binding protein (SREBP), PPAR and other nuclear receptors (Desvergne et al. 2006).
In molecular nutrition, most studies investigate linear gene expression changes in response to treatments or dietary interventions (Dawson et al. 2005). However, it has been shown that many biological interactions between genes and environment (e.g., diet) reportedly occurs in a nonlinear fashion (Chen et al. 2001; Nicholson et al. 2004; Pácha 2000). In the small-intestine, long-chain fatty acids and vitamins undergo concentration-dependent, nonlinear transport across the enterocytes (and such process can even be saturated; see Chen et al. 2001; Pácha 2000). Also,, the relevance of nonlinear responses is considered by the dual intervention point model (Speakman et al. 2011). Such model identifies upper and lower limits at which maximal physiological regulations take place (whereas minimal or no regulations occur within the range of these two limits) and has been proposed to explain the regulation of body weight (and fatness). The upper limit to body fat accumulation in

mice suggests the presence of genetic mechanisms that are activated when a certain limit is attained (Bäckhed et al. 2004; Parks et al. 2013). However, in presence of continual feeding there are some mice strains that display a disrupted set-point mechanism (i.e., fat accumulation is not constrained and continues throughout the whole feeding period; Parks et al. 2013). These studies suggest that the genetic mechanism regulating fat mass gain reacts to the continual feeding with both saturated (logarithm) and unsaturated (exponential) responses. Therefore, the exclusive focus on linear response patterns can represent a limit for the understanding of different biological processes that link food intake to gene expression. An alternative approach that seems particularly suitable for studying gene expression in the small-intestine is to model transcriptomic response to varying concentrations of dietary fat intake using linear and nonlinear fitting, by considering fat intake as the independent variable and gene expression intensity as the dependent one. Nonlinear response types can be used to assess the presence of dose-dependent thresholds (e.g., upper limits to gene expression that can be modelled with the logarithmic function). The presence of upper limit thresholds can be related to excessive fat intake and used to quantify small intestine metabolic capacity.

1.7 Thesis objective

This thesis investigates whether various gene expression response types (i.e., linear and nonlinear) characterize the longitudinal axis of the small intestine as a function of dietary fat doses (i.e., dietary fat intake represents the independent variable and gene expression is the dependent variable). Gene expression is analyzed as a continuous function of fat intake. The aim is categorizing the changes in the response types of differentially expressed genes (i.e., linear response; nonlinear response: logarithmic, exponential, quadratic and cubic) along the small intestine axis. The motivation behind the adoption of these models lies in the idea that excessive fat intake can result in the saturation of some absorptive and

digestive processes. In particular, the linear and exponential responses are associated to processes that are not constrained by any saturation (with the latter referring to an increase of the biological efficiency in presence of high-fat concentration), while the logarithmic response implies a limited capacity of the small intestine to withstand excessive fat intake. This idea is further corroborated by the presence of physiological and anatomical differences from the proximal to the distal section of the small intestine (DeSesso & Jacobson 2001). Such differences are reflected by changes in gene expression and can be related to the following biological mechanisms:

(1) Lipid absorption. Clear differences can be observed for what concerns the absorption capabilities of small intestine, with the middle part known to be particularly active (de Wit et al. 2008). These differences suggest that the efficiency of the most responsive part (i.e., the middle small intestine, in case of absorptive processes) can be saturated by excessive amounts of fat (i.e., logarithmic response).

(2) Lipid transportation. A high-fat diet leads to excess of cholesterol and phospholipids. In the small intestine, this excess concomitantly increases the demand for bile acids that are essential for chylomicron assembly and for a proper fat digestion and absorption. The chylomicron is required for lipid transportation and is accompanied by changes in the transcriptomic levels of genes related to its formation (Desmarchelier et al. 2012). Therefore, modelling gene expression as continuous function of dietary fat intake is relevant to understand whether chylomicron formation can be saturated by excessive fat intake (i.e., logarithmic response) or can promptly respond to it (i.e., either with a linear or exponential response).

(3) Buffer capacity. The development of an obese phenotype occurs beyond the buffer capacity of the small intestine in fat handling (i.e., such buffer capacity has been observed until 20% kcal from fat; see de Wit et al. 2011). Prominent fat-induced differential gene expression was triggered when fat intake exceeded 30% of energy consumption. This

behavior could be due to the fact that fat absorption and transportation are particularly responsive to increasing levels of fat intake (i.e., exponential response), without being associated to comparable changes for what concerns digestive mechanisms (i.e., saturation patterns with logarithmic response).

1.8 Thesis overview

The thesis is organized into five Chapters. Chapter 1 introduces the morphology and physiology of small intestine, and explains how gene expression responses are modulated by dietary fat intake along its axis. It mainly discusses the reasons and motivations behind the approach utilized for the analysis of two different microarray datasets that measure gene expression along the longitudinal axis of small intestine in C57BL/6J mice. This Chapter also explains why the nonlinear modelling of intestinal transcriptomic response to dietary fat intake is particularly relevant for understanding the consequences of high-fat intake.

Chapter 2 describes the methods used for the analysis of the microarray datasets. Two different gene expression datasets were analyzed and are detailed separately in Chapters 3 and 4. Both datasets measured the genome-wide regulation of mRNA abundance in response to dietary lipid and the samples were collected from the intestinal mucosa of C57BL/6J mice.

Chapter 3 reports the results of the analysis performed on the dietary fat responsive gene expression in three sections of the small intestine of C57BL/6J mice. Intestinal transcriptomic data were collected after four weeks of dietary intervention during which mice were fed with different levels of fat (i.e., 10%, 20%, 30% and 45% kcal from fat). Due to the regionalized physiology (i.e., differences in absorption and digestive capability in each of the three sections) and the availability of higher amount of fat in the proximal section, linear and nonlinear (i.e., logarithmic, exponential, quadratic or cubic)

transcriptomic responses to dietary fat intake were tested along the longitudinal axis of the small intestine. The objective was investigating the presence of: (1) gene-specific response types for which a particular gene exhibits the same response type regardless of the intestinal section; (2) section-specific response types, meaning that a particular gene displays unique response types in different intestine sections. We found that: (1) the middle section was the most responsive to dietary fat intake; (2) the relative importance of logarithmic and exponential response was highest in the proximal and distal section, respectively; (3) the majority of the genes were linearly responsive to the fat intake; (4) there was gene-specific, linear response of inflammation related processes in the whole intestine, while the processes related to cholesterol transport and efflux were regulated with exponential response in the middle section.

Chapter 4 is about the analysis of transcriptomic data collected from ten sections of the small intestine of C57BL/6J mice. These mice were fed with three kinds of diet (i.e., high-fat, low-fat or chow) and the dietary intervention lasted 2 weeks. The consumption of dietary fat is essential as it represents a metabolic fuel and plays a key role in the structure of cell membranes. However, when fat is consumed in excess the intestine can adapt to it and increase its absorption (i.e., by increasing the secretion of bile acid for fat digestion). Therefore, three main kinds of transcriptomic response (i.e., linear, logarithmic or quadratic) to dietary fat intake were tested along the longitudinal axis of the small intestine. The main findings were: (1) high-fat intake over-expresses processes related to lipid metabolism and transport, while down-regulating carbohydrate metabolism; (2) middle and distal sections are sensitive to fat whereas the proximal section is responsive irrespective of the diet; (3) transport and metabolic processes were more representative of saturated logarithmic and linear response, respectively.

Chapter 5 provides the conclusions of thesis and describes how our new approach can give new clues to understand the relationships between fat mediated modification of intestinal lipid metabolism and predisposition to obesity.

Chapter 2

Materials and methods

2.1 Transcriptome microarray

This section briefly introduces the fundamentals of genomic expression and the Affymetrix GeneChip technology. The main aim is to describe the transcriptomic datasets that were used to investigate the patterns of gene expression response types (i.e., linear and nonlinear: logarithmic, exponential, quadratic or cubic) as a function of dietary fat gradient along the longitudinal axis of the murine small intestine.

2.1.1 Affymetrix GeneChip technology

Gene expression quantification is mainly performed using Affymetrix GeneChip oligonucleotide microarray and cDNA microarray. The datasets analyzed in this thesis were produced using Affymetrix GeneChip microarrays. These GeneChip arrays were manufactured using spatially patterned, light-directed combinatorial chemical synthesis (Lipshutz et al. 1999). Each array contains up to hundreds of thousands different oligonucleotide probes, which are complementary to the target transcripts. Sets of probe pairs called probesets interrogate particular transcripts. A probe pair consists of a set of two features called perfect match (PM) and mismatch (MM) where the former shows perfect complementarity to the target sequence (cRNA prepared from mRNA) whereas a

single base in the middle of the later shows a mismatch (i.e., the MM probe differs from the PM by a change in the central nucleotide). The benefit for having MM probes and probesets is that they reduce the effects of non-specific hybridization/binding and background signal (Murphy 2002). The intensity of the hybridization between the target sample and the probe array is computed for each probe cell (location for each probe). The image of this probe array is stored as .DAT file. The pixel values per probe cell are calculated and used to produce another (.CEL) file containing the probe intensity. This thesis is based on the analysis of data stored in .CEL files.

2.1.2 Microarray data

This thesis investigates the genome wide effects of dietary lipids in the small-intestine using two distinct microarray transcriptomic datasets. The data were collected from the intestinal mucosa of male C57BL/6J mice. Two experiments were carried out; they were based on two types of dietary intervention. For the first dataset, different doses of dietary fat (i.e., 10%, 20%, 30% or 45% kcal from fat) were supplied. For the second dataset, three kinds of diets (i.e., high-fat, low-fat or chow) were used.

In the first case study, transcriptomic data refer to three sections of the small intestine (i.e., proximal, middle and distal). Mice aged 12 weeks were initially fed with a control diet containing 10% energy intake from fat for 4 weeks. The dietary intervention was carried out for another 4 weeks during which mice were fed with diet containing 10%, 20%, 30% or 45% kcal from fat. Table 2.1 details the composition of the diet. Mice were sacrificed at the end of the intervention. In total, we analyzed 120 samples that represent ten biological replicates for each dietary group in each of the three intestinal sections. The GeneChip arrays used for the analysis were European Nutrigenomics Organization (NuGO) customized Affymetrix arrays (nugomm1a520177mmentrezg); each array contained 16,269 probesets. The microarray data are available at GEO (accession number

GSE26300) and are MIAME (Minimum Information About a Microarray Experiment) complaint. Chapter 3 presents the results obtained by analyzing this dataset.

	10E% fat		**20E% fat**		**30E% fat**		**45E% fat**	
	gm%	***kcal%***	**gm%**	***kcal%***	**gm%**	***kcal%***	**gm%**	***kcal%***
Protein	19	*20*	20	*20*	22	*20*	24	*20*
Carbohydrate	67	*69*	60	*60*	53	*50*	40	*35*
Fat	4	*10*	9	*20*	14	*30*	24	*45*
Other	10	*1*	11	*1*	11	*1*	12	*1*
Total	100	*100*	100	*100*	100	*100*	100	*100*
kcal/gm	3.8		4.1		4.3		4.7	
Ingredient	**gm**	***kcal***	**gm**	***kcal***	**gm**	***kcal***	**gm**	***kcal***
Casein, lactic	200	*800*	200	*800*	200	*800*	200	*800*
L-Cystine	3	*12*	3	*12*	3	*12*	3	*12*
Corn Starch	**427.2**	*1709*	**325.8**	*1303*	**224.3**	*897*	72.8	*291*
Maltodextrin	**100**	*400*	**100**	*400*	**100**	*400*	100	*400*
Sucrose	**172.8**	*691*	**172.8**	*691*	**172.8**	*691*	172.8	*691*
Cellulose, BW200	50	*0*	50	*0*	50	*0*	50	*0*
Soybean Oil	25	*225*	25	*225*	25	*225*	25	*225*
Palm oil	**20**	*180*	**65.2**	*586*	**110.2**	*992*	**177.5**	*1598*
Mineral Mix S10026*	10	*0*	10	*0*	10	*0*	10	*0*
DiCalcium Phosphate	13	*0*	13	*0*	13	*0*	13	*0*
Calcium Carbonate	5.5	*0*	5.5	*0*	5.5	*0*	5.5	*0*
Potassium Citrate, 1 H2O	16.5	*0*	16.5	*0*	16.5	*0*	16.5	*0*
Vitamin Mix V10001*	10	*40*	10	*40*	10	*40*	10	*40*
Choline Bitartrate	2	*0*	2	*0*	2	*0*	2	*0*
Total	**1055**	***4057***	**999**	***4057***	**942**	***4057***	**858**	***4057***

Table 2.1: Diet composition of the intervention used for the case study discussed in Chapter 3. Main ingredients are in bold (e.g., palm oil is the main source of saturated fats whereas corn starch is the main source of carbohydrates).

In the second case study, transcriptomic data were collected from ten sections of the small intestine of male C57BL/6J mice. The microarray platform used for this study was customized Affymetrix mogene11stmmentrezg arrays. Mice received diets containing high-fat (HF), low-fat (LF) or chow. The detailed composition of diet is summarized in Table 2.2. Each diet group contained four biological replicates in the ten sections of the small intestine. Altogether we analyzed 120 arrays and each array consisted of 21,266 probesets. Chapter 4 presents the results of this analysis.

	Chow	**LF**	**HF**
Protein %	**24**	**20**	**20**
Fat%	**6**	**10**	**45**
C16:0	1.2	2.8	18.7
C18:0	0	0.4	1.8
C18:1	2.2	3.2	17.4
C18:2	2.6	3.6	7.1
CHO%	**64**	**70**	**35**
Sugar		27	27
Starch		42	7
Fiber	4	1	1
Nitrogen-free extract	60		
Ash %	**6**		
Total	**100**	**100**	**100**

Table 2.2: Dietary intervention used for the case study presented in Chapter 4. Three types of diet were used: chow, low-fat (LF) and high-fat (HF).

2.2 Background correction and quality assessment

Objective of microarray technology is to quantify the specific hybridization between the target and the probe sequence. However, the measured probe intensities can be confounded

by variations from sources such as technical (non-specific hybridization or optical noise) or manual handling. Therefore, it is necessary to perform preprocessing before obtaining significant biological information. Data preprocessing and quality assessment was performed using the interactive environment provided by the affyPLM package (Bolstad 2004). This package is an updated version of the affy package and is part of BioConductor software project (Gentleman et al. 2004) that is based primarily on the R Statistical Environment (R Core Team 2015).The fitPLM function of the affyPLM package was used to fit probe level linear model to Affymetrix data. This function fits robust Probe Level Linear Models to each probeset and converts Affymetrix GeneChip probe level data (AffyBatch) into a PLMset (Probe Level Linear Models fitted to Affymetrix GeneChip probe level data). The standard series of steps adopted for the preprocessing of raw data are: (1) background adjustment and normalization (to reduce variability within or between arrays), and (2) summarization (a probeset consists of multiple probes and summarization combines these intensities to an expression value). In GeneChip arrays, multiple probes in a probeset represent a given transcript and intensity values are summarized into an expression value.

2.2.1 Background adjustment

Affymetrix introduced MM probe sequences in order to reduce non-specific binding. The default background noise correction involves subtracting MM from PM intensity. However, Irizarry et al. found that this strategy can lead to over-estimation of the expression variance (Irizarry et al. 2003). Therefore, they proposed a new procedure called robust multi-array average (RMA) that corrects the PM intensities whilst neglecting MM. This preprocessing algorithm (RMA) is widely used for Affymetrix microarrays. It includes background correction and also performs quantile normalization and summarization (McCall et al. 2010). Background correction for the dataset presented in

Chapter 4 was implemented using RMA (i.e., it is based on global background adjustment). Using the normalization and the summarization methods of RMA, Wu et al. (2012) proposed another background adjustment algorithm called GCRMA. GCRMA adjusts the intensities via estimators derived from a statistical model that relies on probe sequence information. The dataset analyzed in Chapter 3 was preprocessed using GCRMA. Both RMA and GCRMA produce output data with expression values measured in log 2 base scales. The affyPLM package provides multiple quality assessment tools such as Relative Log Expression (RLE) and Normalized Unscaled Standard Error (NUSE) plots. The Relative Log Expression tool computes RLE values for each probeset in each array by comparing the expression value of a probeset against the median expression value of the probeset across all arrays (Brettschneider et al. 2008). This approach assumes that, in a given array, only a few genes are differentially expressed and therefore RLE values are centered near 0. The Normalized Unscaled Standard Error considers the differences in variability; the standard error estimates for each gene in the PLMset are normalized such that the median standard error for a probeset across all arrays equals 1 (Brettschneider et al. 2008). The quality of the datasets of Chapter 3 was assessed using RLE and NUSE plots.

2.2.2 Variability adjustment and quality assessment

Variations in expression between multiple arrays are categorized into interesting variations and obscuring variations. Interesting variations are contributed by biological differences (e.g., differences between disease and normal tissues), while obscuring variations depend on other factors that are of no biological interest but exist in almost all microarray data (i.e., they can be due to sample preparation, hybridization condition and scanner differences; see Bolstad et al. 2003). Therefore, in order to compare measurements between multiple arrays it is necessary to normalize the obscuring variations, which would otherwise lead to misleading results (Irizarry et al. 2003). Normalization of arrays can be

performed with several methods. There are methods which use information from all arrays; these methods are called complete data methods and include cyclic loess, contrast based method and quantile normalization. Other approaches use baseline arrays such as scaling and non-linear methods (Bolstad et al. 2003). Bolstad et al. (2003) compared the performance of these normalization methods for Affymetrix GeneChip. The comparison showed that the quantile method performs better than the others in terms of speed, variance and bias criteria (Bolstad et al. 2003). Quantile normalization transforms the probe intensities of arrays to have the same distribution. This approach can impose problems to the signals in the tail. However, empirical evidence does not suggest any drawback with this normalization method in practice (Bolstad et al. 2003; Irizarry et al. 2003). Therefore, the usage of this method is preferred over the other methods. This normalization method was applied to the datasets analyzed in in Chapters 3 and 4. Additionally, another normalization technique called UPC (Universal exPression Code) was applied to the dataset presented in Chapter 4. This method uses linear models to correct technological and experimental biases by modelling the genomic base composition and the length of target regions (Piccolo et al. 2013). It assumes that genes of same molecular characteristics constitute same background expression levels and estimates the transcriptional activation level using the mixture model (a specific model used by the function). The mixture model is applied separately for each sample. Furthermore, for the dataset presented in Chapter 4 genes with low variability across samples (usually considered not expressed in cells) are filtered using genefilter library (Bourgon et al. 2010; Gentleman et al. 2012). In general, unexpressed genes are detected as the ones that display low variability across all samples. These genes are filtered and excluded from the study for differential expression to increase the power of the analysis.

2.3 Regression models

Linear regression model was applied to describe the relationship between the dependent/response variable (i.e., gene expression) and the independent/predictor variable (i.e., dietary intervention). The linear relation could also be expressed as:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

Or

$$Gene\ expression\ response_i = \beta_0 + \beta_1 Dietary\ intervention_i + \varepsilon_i$$

Where y_i: The predicted value of the response variable
β_0: The intercept or response value when the predictor variable equals zero
β_1: The slope that quantifies the relation between the predictor variable and the response variable
x_i: The independent or predictor variable
ε_i: The residual error

However, in addition to linear response patterns we also investigated different nonlinear functions. This is because nonlinear processes might be particularly useful to describe fat absorption and digestion in the small intestine (i.e., by modelling how gene expression changes as a function of dietary fat intake when moving along the longitudinal axis of small intestine). Therefore, by combining linear and nonlinear analysis different aspects might be investigated: (1) which is the predominant relationship linking gene expression to dietary fat intake (i.e., linear vs. nonlinear patterns)? (2) Are different biological processes associated to specific response types? (3) How does gene expression change along the longitudinal axis of the small intestine (i.e., gene-specific vs. section-specific responses)? Gene-specific responses are constantly associated to some genes, irrespective of the section where the differential expression occurs. Section-specific responses depend on the section where the genes are differentially expressed (e.g., the same gene can respond either in linear or nonlinear way, and the type of functional response depends on the section where the gene is differentially expressed) and not on the identity of the gene. Linear and

nonlinear gene expression in response to dietary fat intake was modelled in the R Statistical Environment (R Core Team 2015) by using the syntax summarized in Table 2.3.

Response	**Model**	**R syntax**
Linear	$y_i = \beta_0 + \beta_1 x_i$	lm(y~x)
Logarithmic	$y_i = \beta_0 + \beta_1 \log(x_i)$	lm(y~log(x))
Exponential	$y_i = \beta_0 + \beta_1 \exp(x_i)$	lm(y~exp(x))
Quadratic	$y_i = \beta_0 + \beta_1 x + \beta_2 x^2$	lm(y~poly(x,2))
Cubic	$y_i = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3$	lm(y~poly(x, 3))

Table 2.3 R syntax for testing linear and nonlinear responses. Biological interpretations of these models is that the consumption of excessive fat may result in the saturation (i.e., logarithmic) of some biological processes (i.e., absorption or digestion). However, some processes that are not constrained by any saturation may be measured by linear and exponential responses, which describe increased efficiency in presence of highest fat concentration. Quadratic and cubic functions measure processes that follow hyperbolic and oscillating responses, respectively.

2.4 Analysis of differentially expressed genes

In this thesis, the relationship linking gene expression to dietary fat intake along the longitudinal axis of the small intestine was investigated. The transcriptomic data presented in Chapter 3 was fitted with linear and nonlinear (logarithmic, exponential, quadratic and cubic) responses by using dietary fat doses (10%, 20%, 30% and 45% of energy intake) as a continuous predictor of gene expression (i.e., gene expression levels were analyzed as a continuous function of fat intake). Gene expression intensity and fat doses were considered as the dependent and independent variable, respectively. Linear regression was performed on the pre-processed data (i.e., log 2 transformed gene expression); data were pre-processed using the limma library (Smyth 2005). The nonlinear responses were tested by modelling expression as a function of logarithmic, exponential quadratic and cubic transformed fat intake. Multiple test correction of p-values were performed using Benjamini and Hochberg's method (Benjamini & Hochberg 1995) and the significance

threshold for the adjusted p-values was at 0.1. When a gene was significant in more than one response type, the one with the lowest adjusted p-value was selected. In Chapter 4, the transcriptomic data was fitted with linear and nonlinear (logarithmic and quadratic) responses as a function of the interaction between diet and intestinal sections (i.e., both diet and intestinal section are the independent variables; gene expression is the dependent variable). This analysis investigated the transcriptomic response to distinct diets (i.e., high-fat, low-fat, or chow) along the small intestinal axis. The reasons behind the fitting of logarithmic and quadratic responses are: (1) when the amount of fat reaching the intestine is highest, the metabolic processes may display saturated (i.e., logarithmic) response; (2) the digestion process and the microbiota effects are expected to be highest in the most proximal and distal sections, respectively; as such these activities could be effectively modelled with a quadratic response that describes the highest expression in the proximal and distal sections (and lowest in the middle). Diet was represented as ordered categorical values (i.e., chow < low-fat diet < high-fat diet) to illustrate the increasing energy content of fat. The intestinal sections were considered as continuous numerical values (i.e., values from 1 to 10 refer to the longitudinal axis of the small intestine, moving from most proximal to the most distal section). The limma library was used to perform linear regression on the pre-processed data (i.e., log 2 transformed gene expression; see Smyth 2005). For the nonlinear responses, gene expression was modelled as a function of the interaction between diet, and logarithmic or quadratic transformed intestinal sections. Lists of dietary fat responsive (differentially expressed) genes were obtained by making comparisons between samples (i.e., high-fat diet vs. low-fat diet; low-fat diet vs. chow). The p-values were adjusted for multiple test correction using Benjamini and Hochberg's method (1995) and the adjusted p-value threshold of 0.1 was considered. In the case that both linear and nonlinear responses were significant then the one with the lowest adjusted p-value was selected. Various nonlinear functions were associated to specific processes

and have different biological explanations: (1) the logarithmic response reports the presence of saturation mechanisms; (2) the exponential response corresponds to extremely efficient processes in the presence of high-fat intake; (3) the quadratic response refers to enzyme kinetics; (4) the cubic response is expected to be associated to circadian rhythms with oscillating trends. The limma library computes moderated t-statistics for each probe (and for each contrast). It differs from the ordinary t-statistics because: (1) the standard errors across genes are moderated to a common value by using the simple Bayesian model (i.e., the standard error measured from the estimated log-fold changes are moderated towards a common value by borrowing information across genes); (2) the degrees of freedom are increased to be consistent with the smoothened standard errors (Smyth 2005).

2.5 Correlation network analysis

The correlation patterns of the diet responsive differentially expressed genes were analyzed by constructing a co-expression network (i.e., each node represents the gene expression profile and the pairwise correlation between expressions are used to generate the edges between genes) using the WGCNA library (Langfelder & Horvath 2008). The hierarchical clustering method implemented in the WGCNA library was used to identify clusters of highly interconnected genes (in graph theory, these clusters are also called modules) in the dataset analyzed in Chapter 4 (Kaufman & Rousseeuw 1990; Langfelder & Horvath 2008). The topological overlap measure of the expression data was inferred from the adjacency matrix (i.e., calculated from the co-expression similarity by using the principle of scale-free topology) and this was used to define the interconnectedness of clusters. In order to avoid identifying clusters with similar expression profiles, the co-expression similarities of all clusters were further evaluated by grouping the correlated clusters on the basis of their eigengene profiles (i.e., the most characteristic gene expression profile is considered as the representative of the module - it is the first principal component of the module; see

Langfelder & Horvath 2008). Most biologically interesting clusters (i.e., clusters that showed significant correlation with the ten intestinal sections) were obtained by finding the most significant association between the eigengene profiles of each cluster with the ten intestinal sections (Langfelder & Horvath 2008).

2.6 Functional annotation of the differentially expressed genes

The functional annotation of the differentially expressed genes found with both datasets (i.e., in Chapters 3 and 4) were performed with two different libraries in R. The over-represented Gene Ontology Biological Processes (GOBP) relative to each section (and for each response type) in the first dataset (i.e., Chapter 3) were investigated using the hypergeometric test. This test was performed using HTSanalyzeR library, which tests for the overlap between hits and gene sets (such as Gene Ontology, MSigDB, and KEGG; see Wang et al. 2011). The p-values were adjusted using Benjamini and Hochberg's method and the significance threshold was 0.1 (Benjamini & Hochberg 1995). Each significant cluster identified with correlation network analysis (i.e., the second dataset; see Chapter 4) was tested for enriched GOBP terms using the gage library (Luo et al. 2009). Three kinds of GOBP enrichment tests were performed for each cluster: (1) by ignoring the dietary effect, the processes that were enriched along the longitudinal axis of the small intestine were obtained; (2) high-fat diet responsive processes were investigated by considering high-fat diet samples as the target and the low-fat diet samples as the reference; (3) low-fat diet responsive processes were investigated by taking low-fat diet samples as the target and chow diet samples as the reference. For each analysis, three kinds of outputs were obtained based on the regulatory direction of the GOBP: (1) up-regulated processes; (2) down-regulated processes; (3) processes which include gene regulation in both directions, such as canonical signaling pathways (Luo et al. 2009). The Benjamini and Hochberg's method was used to obtain adjusted p-values and the significance threshold was at 0.1. For each

cluster, the enriched GOBP terms were grouped based upon GO similarities (i.e., by considering information content measure of GO terms within a specific ontology - biological process in our case) as implemented in GOSim library (Fröhlich et al. 2007). The information content values for each GO term are precomputed. These values are based on the number of times a specific GO and its gene products or neighboring (i.e., direct or indirect) offspring are annotated into the GO. In order to test the nonlinear responses in each network cluster the over-represented GO terms that refer to metabolic and transport processes were selected for further analysis. Also, the clusters that showed the highest significance to diet-responsive GOBP and that displayed the highest correlation with the ten intestinal sections (i.e., as computed using WGCNA library) were considered for the subsequent analysis (i.e., to find the predominant transcriptomic responses for each enriched process and examine which intestinal segments are most responsive to diet). Each of these clusters was characterized by the GO term with the highest information content value. This selection was made by identifying more specific GOBP terms (e.g., lipid metabolic process prevails over general terms like metabolic process).

2.7 Biological process trend analysis

In order to check if specific response types (i.e., linear, logarithmic, exponential, quadratic or cubic) are representative of certain biological processes or change their relative importance along the small intestine axis, GOBP trends were analyzed from the most proximal to the most distal section (results of such analysis are presented in the Chapter 3). In the Chapter 3 we aimed at finding: (1) the GOBP terms that were composed of genes that preserved their response type in all intestine sections (i.e., proximal: linear – middle: linear – distal: linear); (2) the GOBP terms that included genes that changed the shape of their response type along the different small intestine sections (e.g., proximal: linear – middle: logarithmic; middle: logarithmic – distal: linear; middle: linear – distal:

exponential). Two additional analyses were performed in Chapter 4 to identify dietary mediated predominant response types in each of the ten sections. First, the most characteristic (gene-level) response pattern of each GOBP per cluster was considered as the dominant one (i.e., the linear or nonlinear response type for which the highest gene count was recorded). Second, for each gene the intestinal sections with the highest frequency of high/low expression values were considered to be most responsive to diet.

Chapter 3

Nonlinear transcriptomic response to dietary fat intake in the small intestine of C57BL/6J mice[1]

3.1 Background

Overconsumption of food that are rich in saturated fats leads to excessive energy intake and is strongly linked to metabolic disorders such as obesity, diabetes, cardiovascular diseases and some forms of cancer (Balwierz et al. 2009; Cheung & Mao 2012; Desmarchelier et al. 2012; Walley et al. 2006; Zhang et al. 2009). As the primary source of dietary fat uptake, the small intestine plays a key role in governing nutritional health (Kaminsky & Zhang 2003; Kondo et al. 2006; Petit et al. 2007; Thomson et al. 1986; de Wit et al. 2008). The intestinal absorptive capacity is enhanced by numerous fingers like projections of the mucosal membrane called villi, and there exist regionalized anatomic and physiological differences from proximal duodenum to distal ileum (Figure 1.1). The middle section has longer villi in comparison with the proximal and distal sections. It covers nearly half of the intestinal length and is characterized by the highest absorptive

[1] Chapter 3 includes the findings of the following manuscript: Nyima, T. et al., Nonlinear transcriptomic response to dietary fat intake in the small intestine of C57BL/6J mice. Under revision in *BMC Genomics*.

capacity. The distal section contains shorter villi and less absorptive capacity (Desmarchelier et al. 2012; Goncalves et al. 2015).

The small intestine acts as a gatekeeper between the diet and the body and can directly metabolize or block the uptake. Recent studies have demonstrated a strong intestinal transcriptomic response to dietary fat intake. Kondo et al. (2006) found high fat-induced up-regulation of lipid metabolism-related genes (e.g., *Mod1*, *Cyp4a10*, *Acot1* and *Acot2*) in the small intestine of C57BL/6J mice, with negligible effects observed in the liver, muscle and white adipose tissue. Fat intake triggers the down-regulation of ABC half-transporters (*Abcg5* and *Abcg8*) in liver and intestine thus leading to increased levels of sterols in diabetic rats (Bloks et al. 2004). Biological processes like inflammatory response and cell cycle were highly up-regulated in the small intestine of C57BL/6J mice during dietary fat induced development of obesity and insulin resistance (de Wit et al. 2008). The absorption capacity of intestine displays adaptability in response to dietary fat composition such as enhanced intestinal cell proliferation, synchronization of fatty acid uptake and lipoprotein secretion, and altered transport processes (Petit et al. 2007; Thomson et al. 1986).Section-wise studies focused on gene-specific responses: (I) Simon et al. (2011) found increased distal gut hormone response to a high fat diet in apoA-IV knockout mice; (II) Nassir et al. (2007) observed sharp decreasing gradient in CD36 levels from proximal to distal intestine. de Wit et al. (2011) showed prominent effect of dietary-fat doses on gene expression, mainly in the proximal and middle sections. They concluded that differentially expressed genes correlated with the development of obesity, and the main shift towards an obese phenotype was observed when an amount of energy included between 20% and 30% was derived from fat. Previous research studies investigated the consequences of fat intake on transcriptomic response of intestine by individual, pairwise comparisons between control (i.e., baseline) diets and treatments (e.g., high fat vs. low fat; see Kondo et al. 2006; Westergaard & Dietschy 1974; de Wit et al. 2008). Their goal was detecting a significant

difference between control and treatments, rather than quantifying the strength of the response as a function of different levels of fat intake. Our work is a re-analysis of the data produced by (de Wit et al. 2011). We propose a novel approach to study whether linear or nonlinear response types characterize gene expression in the small intestine (i.e., dietary fat intake represents the independent variable and gene expression level is the dependent variable). We investigate dose-dependent transcriptomic response to dietary fat along the horizontal axis (proximal, middle and distal sections) of the small intestine of C57BL/6J mice. Mice were fed with 10%, 20%, 30% or 45% of energy (E%) derived from fat for four weeks (n = 10 mice/diet) and corresponding gene expression levels served to fit dose-dependent responses. Due to the unique morphological and functional characteristics of each intestinal section (Desmarchelier et al. 2012; Hollander & Truscott 1976; Shaw-Smith & Walters 1997; Thomson et al. 1986; Westergaard & Dietschy 1974; Westergaard & Dietschy 1976), we expect to observe variations in transcriptomic response from proximal to distal section (Figure 1.1a). Our focus is mainly upon evaluating possible nonlinear relationships between gene expression and dietary fat percentage in the three sections. We aim at categorizing changes in the response types of differentially expressed genes (i.e., linear response; nonlinear responses: logarithmic, exponential, quadratic or cubic) among each of the three sections of the small intestine (Figure 3.1). The main advantage of studying continuous transcriptomic responses as a function of dietary fat intake is related to the chance of combining the information of different treatments into a unique picture. This helps in better understanding how the response to highest levels of fat intake is attained. Thus, we can investigate whether biological processes are overloaded by highest levels of fat content (i.e., logarithmic response type) or modulated in order to cope with them (i.e., exponential response type). The presence of a constant, linear response type along the whole small intestine can indicate spatial-independent mechanisms (i.e., this applies to processes that respond in a gene-specific manner, independently of the intestine

section where they occur). For example, such an approach can be used to illustrate whether an obese phenotype (I) corresponds to a 'tipping point' of the system (i.e., exponential response vs. linear and logarithmic responses) or (II) is limited by the metabolic capacity of the system (i.e., logarithmic response vs. linear and exponential responses). Moreover, distinguishing among the shape of the response types in differentially expressed genes adds a further qualitative level to the description of the biological processes (i.e., in addition to the magnitude and direction of the regulation). The broadest goal of our study was investigating whether the response types are gene-specific (i.e., a given gene always exhibits the same response type, regardless of intestinal section) or intestinal section-specific (i.e., for the same gene, the response type displays unique patterns in the three intestine sections).

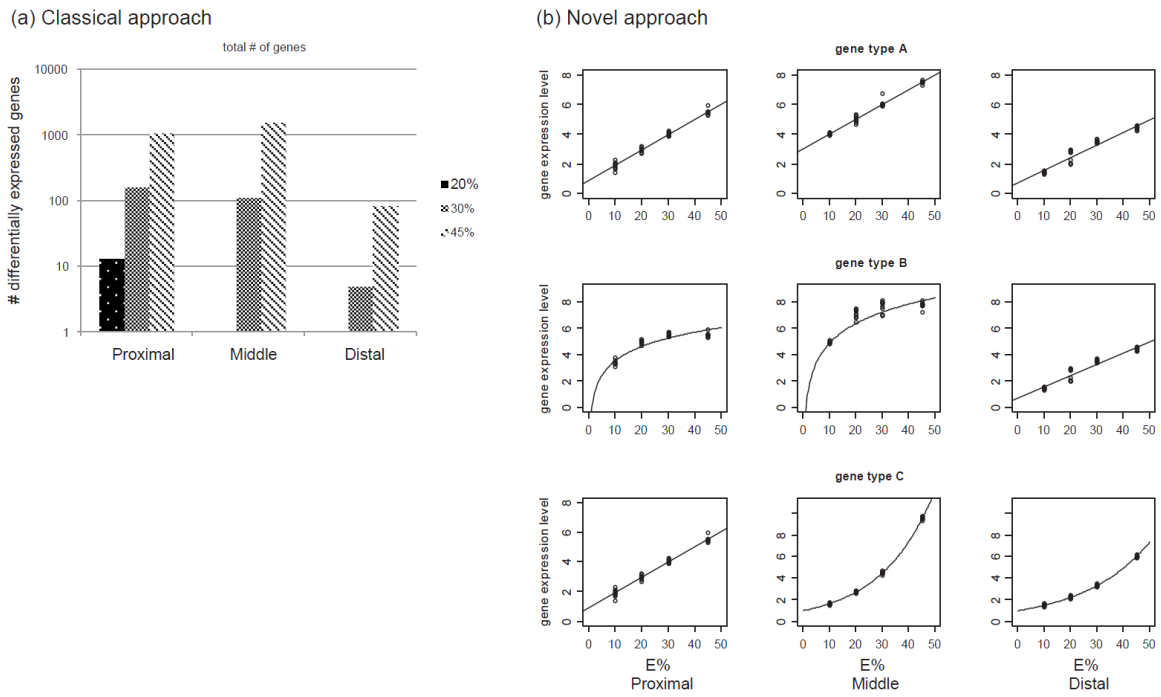

Figure 3.1: Comparison between classical and novel approach for identifying differentially expressed genes. We use as a reference the case study of mice fed with four levels of fat representing 10%, 20%, 30% and 45% of total energy intake; gene expression is measured in three sections of small intestine: (a) in the classical approach, the total number of differentially expressed genes is determined by comparing the control diet (10 E%) with treatments (20 E%, 30 E% and 45 E%); (b) in the novel approach, all expression levels of each gene (in the three sections) are considered to fit linear or nonlinear (i.e., logarithmic and exponential, in this case) response types. Three hypothetical patterns are described in (b): (I) gene type A preserves a linear response in the three sections of small intestine; (II) gene type B displays saturation (i.e., logarithmic response) in the proximal and middle sections; (III) gene type C shows exponential responses in most downstream sections (i.e., middle and distal parts of small intestine). Transcriptomic response of gene type A is gene-specific while the patterns observed for gene type B and C are section-specific.

3.2 Materials and methods

3.2.1 Ethics statement

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

3.2.2 Dietary intervention

At twelve weeks of age, mice were fed for an initial period of four weeks with a control diet containing 10% of energy intake from fat. The main reason for this relatively long run-in period was that we wanted to be sure that the effects of chow diet were highly diluted/not present when starting the dietary intervention. In addition, since we wanted to investigate the effect of diet-induced obesity in adults, we choose to start the experiment when the mice were 16 weeks old (mice aged 12 weeks are considered young adults that are still growing). After the initial period, mice were divided into four groups that received 10%, 20%, 30% or 45% kcal from fat (see Table 2.1 for dietary composition). The dietary intervention lasted four weeks and then mice were killed by cervical dislocation after five hours fasting and anesthesia with 1.5% isoflurane. The isoflurane was evaporated in a vaporizer using a mixture consisting of 70% nitrous oxide and 30% oxygen. Ten biological replicates were carried out for each diet group in the three sections of the small intestine (i.e., altogether we analyzed 120 samples).

3.2.3 Microarray data

We analyzed microarray transcriptomic data from the intestinal mucosa of male C57BL/6J mice. The small intestine was divided in three equal parts: proximal, middle and distal

sections. These three parts were chosen because of practical reasons; when dissecting the small intestine one has to process the tissue quickly to avoid RNA degradation. The 1st part (i.e., proximal section) consists of the duodenum plus proximal jejunum, the 2nd part (i.e., middle section) corresponds to jejunum, and the 3rd part (i.e., distal section) includes the distal jejunum and ileum. Detailed protocols on dietary intervention and RNA extraction are described by de Wit et al (de Wit et al. 2011). The microarray platform used for this study is nugomm1a520177mmentrezg, a custom Affymetrix mouse array containing 16,269 probesets. The NuGO arrays are custom designed Affymetrix GeneChip arrays, designed by the European Nutrigenomics Organisation (NuGO) and manufactured by Affymetrix. These arrays contain in part common probe sets that are also present on standard Affymetrix arrays and in part newly designed probe sets (GEO platform GPL7440). The microarray data used for our analyses are MIAME compliant, available at GEO (accession number GSE26300). Data pre-processing and quality assessment, statistical analysis to identify differentially expressed genes and pathway analysis have been carried out in the R Statistical Environment (R Core Team 2015).

3.2.4 Data pre-processing and quality assessment

We used the affyPLM (PLM = Probe Level Model) library for data preprocessing and quality assessment. We applied the fitPLM function that fits iterative reweighted least square M-estimation regression to the probe intensity (Bolstad 2004). Background intensities (optical noise and non-specific binding) were adjusted with the GCRMA library (Wu et al. 2007). Such adjustment is obtained via estimators derived from a statistical model that uses probe sequence information. The GCRMA library has been shown to perform particularly well in adjusting background intensity in Affymetrix Genechips (Kroll et al. 2008). After background adjustment, technical variability between arrays was adjusted by quantile normalization (Bolstad 2001). The quality of the PLMset object was

assessed by plotting Relative Log Expression (RLE) and Normalized Unscaled Standard Error (NUSE; Brettschneider et al. 2008). Genes with low variability across samples are usually considered as not expressed. This is motivated by the observation that, in general, unexpressed genes are detected most reliably through low variability of their features across samples. Non-specific filtering of the genes was made with the genefilter library (Bourgon et al. 2010; Gentleman et al. 2012). The pOverA R function was used for variance-based filtering; genes with unlogged intensity above five, in at least five arrays, were chosen for the subsequent analysis. Genes without Entrez Gene ID and Affymetrix quality control probe-sets were excluded. After data pre-processing and quality assessment we selected 14,952 genes in the proximal section, 14,933 in the middle and 14,925 in the distal.

3.2.5 Statistical analysis to identify differentially expressed genes

We investigated the statistical relationship between gene expression and dietary fat intake. We fitted linear and nonlinear (i.e., logarithmic, exponential, quadratic and cubic) responses describing gene expression levels (dependent variable) as a function of dietary fat intake (independent variable). We considered nonlinear responses that reflect: (I) blunted differential expression at higher fat intake, which may indicate overloading of the relevant biological process (logarithmic curve); (II) progressively stronger differential expression with increasing fat intake (exponential curve); (III) parabola-like differential expression (quadratic function); (IV) oscillating trends (cubic function). The array data were log2 transformed (GCRMA normalized data). All responses (i.e., linear and nonlinear) were tested by modelling log2 transformed expression as a function of fat intake (i.e., with fat providing 10, 20, 30 or 45 E%). When comparing the diets of 40 mice (n = 10 mice per diet group), there was no significant difference with respect to total food intake in grams per day (see Table 3.1). However, caloric intake increased with increasing fat

percentage (and a linear relationship existed between the % of kcal from fat intake and the actual grams of fat intake - i.e., soybean oil and palm oil). Fat intake (the independent variable) is considered as continuous predictor of log2 transformed gene expression. We used the limma library (Smyth 2005) to perform linear regression on GCRMA normalized data (i.e., log2 transformed gene expression), and tested nonlinear responses by modeling expression as a function of logarithmic-, exponential-, quadratic- and cubic-transformed fat intake. To identify differentially expressed genes we performed multiple testing correction using Benjamini and Hochberg's false discovery rate (1995; FDR, with 0.1 significance threshold). It should be noticed that many studies regularly adopt an adjusted p-value threshold of 0.1 for identifying differentially expressed genes (e.g., Daves et al. 2011; Stevens et al. 2013). In the case that both linear and nonlinear responses were significant for a given gene, we selected the one with the lowest p-value. The robustness of our analysis was then tested by comparing the results obtained with two criteria of model selection (i.e., using p-values and Akaike's information criterion - AIC).

Feed intake in grams per day (mean and SEM, n = 10)							
10E% fat		20E% fat		30E% fat		45E% fat	
MEAN	SEM	MEAN	SEM	MEAN	SEM	MEAN	SEM
3.396	0.060	3.279	0.072	3.404	0.034	3.196	0.073
Caloric content feed (kcal/gm)							
3.8		4.1		4.3		4.7	
Caloric intake (kcal per mouse per day); mean (n = 10)							
12.905		13.443		14.635		15.020	

Table 3.1 Food and caloric intake in week 4. Data refer to the week the mice were killed. We summarized the mean and standard error of the mean (SEM), based on n = 10 per diet group. The caloric intake consistently increased from 10 to 45% diets, but the absolute amounts of food intake varied in a non-significant way.

3.2.6 Functional analysis

We investigated whether fat-responsive genes were enriched within distinct biological processes (defined as Gene Ontology Biological Process - GOBP), and tested whether each intestinal section displayed unique transcriptomic response to fat intake. We performed hypergeometric tests to functionally characterize groups of differentially expressed genes. We investigated over-represented Gene Ontology Biological Process (GOBP) in each section (adjusted p-value < 0.1; such threshold is commonly used for gene set enrichment analysis - e.g., Harris et al. 2009; Raymond et al. 2010) using the library HTSanalyzeR (Wang et al. 2011). To understand whether specific response patterns (i.e., linear, logarithmic, exponential, quadratic or cubic) characterize certain biological processes, or change according to the intestinal sections, we assessed trends of GOBP terms passing from the proximal to the distal region. We studied GOBP terms that preserved the same response pattern in different intestinal sections (e.g., proximal: linear - middle: linear - distal: linear), and analyzed those changing response pattern between the sections (e.g., proximal: linear - middle: logarithmic; middle: logarithmic - distal: linear; middle: linear - distal: exponential). A flowchart of the overall analysis is shown in Figure 3.2.

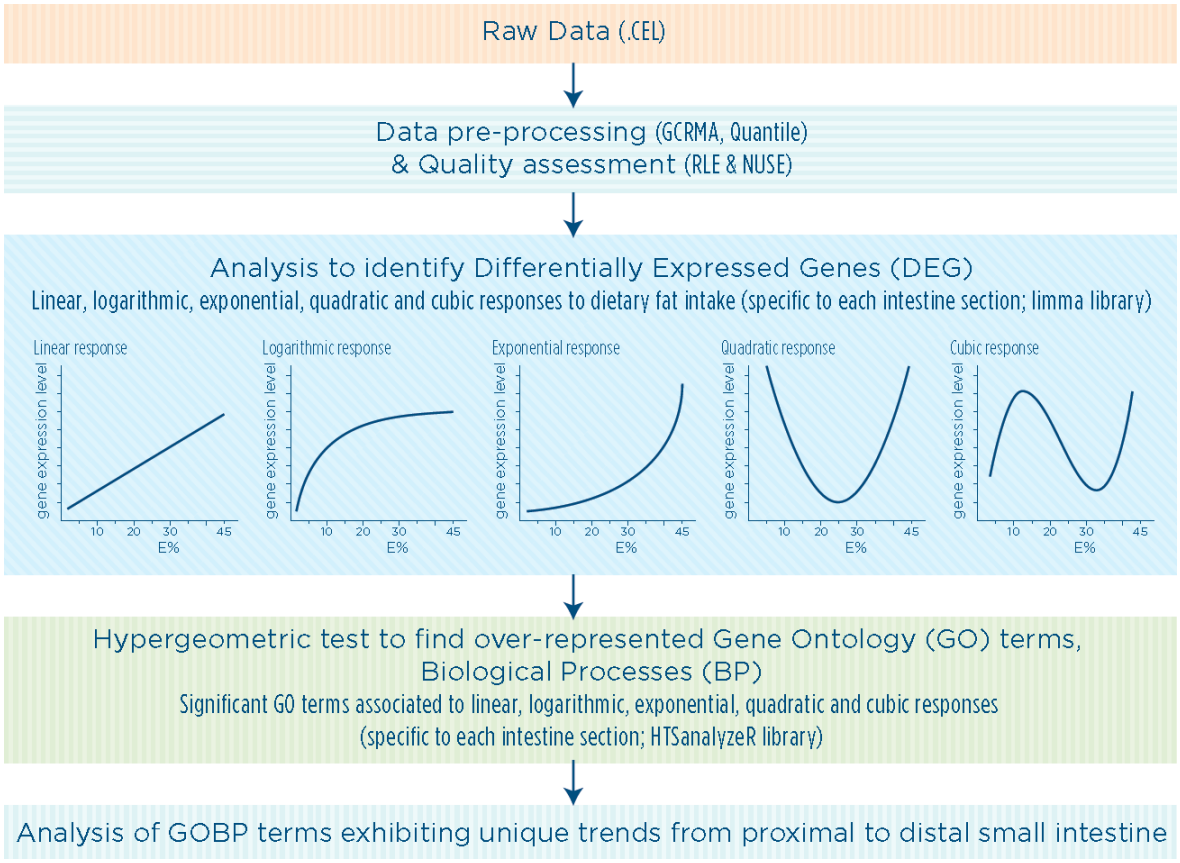

Figure 3.2 Schematic representation of workflow adopted for the analysis of microarray data. We identified linear and nonlinear response patterns for significantly over-represented GOBP terms. Trends for the response patterns specific to each term were examined from proximal to distal section.

3.3 Results

3.3.1 Linear and nonlinear gene expression

The count of genes that respond in a significant way (either linear or nonlinear) to fat intake varies between the three sections (Table 3.2).

Response (Gene expression vs. Fat intake)	**Proximal**			**Middle**			**Distal**		
	Gene (count)	Relative %	GOBP (count)	Gene (count)	Relative %	GOBP (count)	Gene (count)	Relative %	GOBP (count)
Linear	**1053**	**53.29**	25	**1410**	**56.65**	27	**185**	**59.11**	26
Logarithmic	**678**	**34.31**	1	**458**	**18.40**	22	**26**	**8.31**	-
Exponential	**139**	**7.03**	-	**499**	**20.05**	9	**84**	**26.84**	8
Quadratic	**61**	**3.09**	-	**87**	**3.50**	7	**9**	**2.88**	-
Cubic	**45**	**2.28**	-	**35**	**1.41**	2	**9**	**2.88**	-

Table 3.2: Count of genes that respond in a significant, dose-dependent way to fat intake. Results were extracted using the limma library and are specific to each section of small intestine; adjusted p-value < 0.1. Genes are classified based on the mathematical function used to fit their expression in response to fat intake (linear, logarithmic, exponential, quadratic or cubic). For each section, we summarize the relative percentage of genes characterized by the five response types (%) and the number of significant GO terms (Biological Processes - GOBP) found with hypergeometric test (adjusted p-value < 0. 1). Only GOBP terms with, at least, six differentially expressed genes (in the whole small intestine) have been taken into account.

The highest number of genes showing linear, exponential or quadratic transcriptomic response to fat intake was found in the middle section, but the highest number of genes displaying either logarithmic or cubic patterns of expression in response to fat intake was in the proximal section. The lowest number of differentially expressed genes, for all five

response types, was observed in the distal section. The proportion of differentially expressed genes exhibiting a significant linear relationship with the dietary fat intake was always above 53%. In the proximal and distal sections, the relative importance of logarithmic and exponential response, respectively, was the highest. Both quadratic and cubic response types are of marginal importance in all intestine sections (hereafter, given the fact that quadratic and cubic responses are associated with around 5% of differentially expressed genes, the focus will be on linear, logarithmic and exponential patterns only). The number of genes displaying a logarithmic response to fat intake dropped from the proximal to the distal section (the percentage decreases from 34.31% in the proximal to 8.31% in the distal section). The opposite trend is found for the exponential response: the percentage increases from 7.03% in the proximal to 26.84% in the distal section. Thus, although the highest number of genes responding in an exponential way to fat intake was in the middle section, an increasing relative importance of this response is observed from the proximal to the distal part of the small intestine (Figure 3.3). In the three sections, if we consider the changes in gene expression that were significantly associated with varying levels of fat intake we observed a prevalence of up-regulated genes (Figure 3.4). These results were not affected by the model selection criteria chosen for the analysis (the trends of differentially expressed genes identified using AIC are summarized in the Table 3.3.

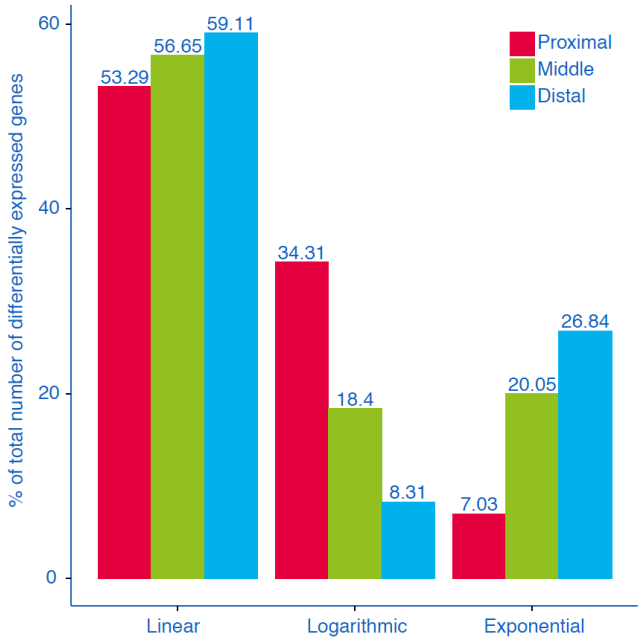

Figure 3.3: Percentages of differentially expressed genes that respond in a linear, logarithmic and exponential way to fat intake. Histograms are grouped according to response type, while different colors are associated to three small intestine sections. Linear response is highly represented in all sections, while the relative importance of logarithmic (exponential) response decreases (increases) from the proximal to the distal part of small intestine.

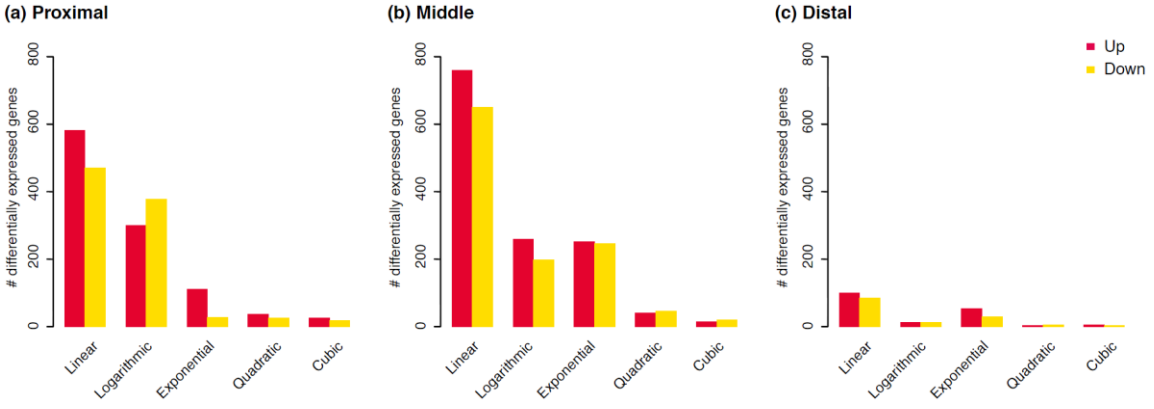

Figure 3.4: Differentially expressed genes that are up- or down-regulated in response to increasing fat intake. Results were obtained using the limma library, modeling gene expression as a function of raw fat intake data (i.e., linear model), logarithm-, exponential-, quadratic- or cubic-transformed fat intake values; adjusted p-value < 0.1. In each plot, y-axis represents the gene count and x-axis corresponds to response types (linear, logarithmic, exponential, quadratic and cubic).

Response (Gene expression vs. Fat intake)	**Proximal**			**Middle**			**Distal**		
	Gene (count)	Relative %	GOBP (count)	Gene (count)	Relative %	GOBP (count)	Gene (count)	Relative %	GOBP (count)
Linear	**1219**	**48.18**	6	**1439**	**50.03**	16	**366**	**50.76**	3
Logarithmic	**719**	**28.42**	7	**443**	**15.40**	11	**45**	**6.24**	-
Exponential	**148**	**5.85**	2	**461**	**16.03**	17	**121**	**16.78**	6
Quadratic	**212**	**8.38**	-	**326**	**11.34**	-	**85**	**11.79**	-
Cubic	**232**	**9.17**	1	**207**	**7.20**	-	**104**	**14.42**	4

Table 3.3 Count of genes that responded in a significant, dose-dependent way to fat intake (model selection based on AIC). We identified the significant responses in three sections of small intestine (limma library, adjusted p-value < 0.1). Differential gene expression was modelled as either a linear or nonlinear (i.e., logarithmic, exponential, quadratic or cubic) function of dietary fat intake. For each section, we reported the percentage of genes per response type (%) and the number of significant GO terms (Biological Processes - GOBP; hypergeometric test, adjusted p-value < 0. 1). Only GOBP terms with, at least, six differentially expressed genes (in the whole small intestine) have been taken into account. Outcomes obtained by using AIC for model selection were coherent with the model selection based on the lowest p-value (see Table 3.2). In particular, with AIC we found: (I) the prevalence of linear-responding genes in all intestine sections; (II) the decreasing relative importance of the logarithmic response when moving from the proximal to the distal section (while the opposite pattern holds for the exponential response); (III) the marginal relevance of other response types (i.e., quadratic and cubic functions).

3.3.2 Functional annotation

In the previous section, we classified the genes according to their category of response to fat intake (i.e., using linear, logarithmic, exponential, quadratic and cubic response types). We found two main classes: (I) genes that preserve linear responses along the whole longitudinal axis of small intestine; and (II) genes that alter the shape of their

transcriptomic response, depending on intestinal section. Due to the progressive absorption that occurs in the small intestine the amount of fat decreases from the proximal to the distal section. This pattern correlates well with the decrease of the relative importance of overload/logarithmic response, and with the increase in relative importance of exponential response. The broadest objective of our study was to characterize the biological processes that are associated with: (I) genes that maintain linear response types (gene-specific processes), and (II) genes that change their transcriptomic response as a function of fat availability (section-specific processes). To this aim we applied hypergeometric test to investigate the presence of GOBP terms that are significantly enriched with fat-responsive genes. We analyzed lists of genes associated with different response patterns to fat intake (i.e., linear, logarithmic or exponential). First, we considered full lists of genes, including both up- and down-regulated ones. Second, we performed gene set enrichment analysis by taking into account the direction of change (i.e., using either up- or down-regulated genes only). We found terms that are over-represented in more sections and, in some cases, they are characterized by changes in the response type along the longitudinal axis of small intestine (section-specific processes). This served to understand whether GOBP terms in different sections preserve their response type (e.g., APR, acute-phase response exhibited a linear response type in all three sections, including both up- and down-regulated genes, and can be classified as a gene-specific process: thus, fat intake triggered the same response pattern, independently of intestinal sections; Tables 3.4 – 3.6) or showed some changes (e.g., cholesterol homeostasis was characterized by linear response in the proximal section and changed to logarithmic in the middle; mostly up-regulated genes, Tables 3.4 – 3.6). The complete list of genes associated with specific processes (i.e., we do not include details on three generic GOBP terms: metabolic process - GO:0008152; oxidation-reduction process - GO:0055114; transport - GO:0006810) that displayed unique patterns along different sections of the small intestine is summarized in Table 3.6 (details on up-

and down-regulated genes are available). The Tables 3.7 – 3.10 list GOBP terms and genes with characteristic response types along the small intestine; they refer to strictly up- and down-regulated processes, respectively. We identified distinctive patterns for inflammation-related pathways (i.e., acute-phase response and negative regulation of caspase activity) and cholesterol-related processes (i.e., cholesterol transport, cholesterol homeostasis and cholesterol efflux).

(I) Acute-phase response (GO:0006953): genes belonging to this GOBP term showed a constantly linear response in all three sections of small intestine, with mixed direction of regulation (i.e., both up- and down-regulated genes; see Table 3.6). The only gene differentially expressed in all three sections was *Serpina1b* (up-regulated).

(II) Negative regulation of caspase activity (GO:0043154): caspase activity plays an essential role in apoptosis and inflammation. This is a section-specific process: from middle to distal section we observed a change from linear to exponential expression (Table 3.4). The genes *Birc5* (up-regulated) and *Igbp1* (down-regulated) were significantly responsive in the middle and the distal sections. *Prdx3* and *Gpx1* were differentially expressed and up-regulated in the middle section only.

(III) Cholesterol transport (GO:0030301): the response of differentially expressed genes related to this GOBP term changed from linear to exponential when passing from the proximal to the middle section (Table 3.4). The gene *Abca1* was down-regulated in both proximal and middle small intestine. *Cd36* showed up-regulation proximally, while all differentially expressed genes that showed an exponential response in the middle section were down-regulated (e.g., *Abcg1* and *Scarb1*).

(IV) Cholesterol homeostasis (GO:0042632): up-regulated genes involved in this process exhibited overload in the middle section only (i.e., they were linearly significant proximally and in the distal section, but displayed logarithmic response in the middle; Tables 3.7 – 3.8). Most of the genes involved in cholesterol homeostasis were up-regulated

(Tables 3.6 – 3.8). For example, the gene *Pla2g10* was up-regulated in both proximal and middle section, and the same positive regulation was displayed by *Apoa4* when moving from the middle to the distal part of small intestine (Table 3.8). These findings may illustrate how excessive fat intake can have detrimental consequences (i.e., logarithmic response with an upper limit to gene expression) in processes involved in the maintenance of the cholesterol steady state in cells, especially in the middle section.

(V) Cholesterol efflux (GO:0033344): this GOBP term shares many genes with cholesterol transport and homeostasis, but exhibits a less sharply defined behavior. Up-regulated genes showed logarithmic and linear response in the middle and distal sections, respectively (e.g., *Apoa4*). Down-regulated genes displayed linear and exponential response in proximal and middle parts, respectively (e.g., *Abca1*). The overload of the process that regulates steady state of cholesterol within cells (i.e., cholesterol homeostasis) represents a bottleneck, and the negative, exponential response found for cholesterol transport should be considered in relation to it (with the sharp down-regulation that is triggered by the highest level of fat intake only; i.e., 45E%). It seems that the directed movement of cholesterol into or between cells is impaired when cholesterol level reaches its carrying capacity.

In summary, by completing the analysis of transcriptomic response patterns (i.e., linear and nonlinear response types) with functional annotation we highlighted three main mechanisms of action in the small intestine: (I) acute-phase response (an inflammatory-related process) is not section-specific and exhibits a linear regulation along the whole small intestine; (II) lipid absorption and transportation are particularly active in the middle section, but the coupling with other overloaded functions can limit the processing capacity (see the logarithmic, up-regulation of cholesterol homeostasis and associated exponential, down-regulation of cholesterol transport); (III) in the distal section, an exponential

response characterizes the interplay between up- and down-regulated genes involved in the negative regulation of caspase activity, likely having a role in apoptosis and inflammation.

Section and Responses	GO ID	GOBP	Proximal			Middle			Distal		
			Set Size	Hits	Adjusted p-value	Set Size	Hits	Adjusted p-value	Set Size	Hits	Adjusted p-value
Prox – Mid – Dist (all linear)	GO:0006953	acute-phase response	22	5	0.09	22	6	0.08	21	3	0.03
Prox – Mid (lm - lm)	GO:0006629	lipid metabolic process	184	33	<0.01	183	46	<0.01			
	GO:0006631	fatty acid metabolic process	68	19	<0.01	67	28	<0.01			
	GO:0006635	fatty acid beta-oxidation	21	9	<0.01	21	11	<0.01			
	GO:0006637	acyl-CoA metabolic process	19	7	<0.01	19	6	0.04			
	GO:0007040	lysosome organization	18	5	0.03	18	6	0.04			
	GO:0015031	protein transport	386	45	0.01	387	60	<0.01			
	GO:0022900	electron transport chain	69	14	<0.01	69	18	<0.01			
	GO:0055085	transmembrane transport	390	42	0.07	391	57	0.02			
Prox – Mid (lm - log)	GO:0042632	cholesterol homeostasis	33	11	<0.01	33	4	0.10			
Prox – Mid (lm - exp)	GO:0030301	cholesterol transport	15	7	<0.01	15	4	0.03			
Mid – Dist (lm - lm)	GO:0006520	cellular amino acid metabolic process				16	5	0.07	16	2	0.05
	GO:0016042	lipid catabolic process				69	14	0.05	69	5	0.03
	GO:0051262	protein tetramerization				17	5	0.09	17	2	0.05
Mid – Dist (lm - exp)	GO:0043154	negative regulation of caspase activity				36	10	0.02	36	2	0.06
Mid – Dist (log - lm)	GO:0006644	phospholipid metabolic process				17	4	0.01	17	2	0.05
	GO:0006749	glutathione metabolic process				24	4	0.04	24	2	0.08
	GO:0045859	regulation of protein kinase activity				16	4	0.01	16	2	0.05

Table 3.4: Over-represented Gene Ontology Biological Process (GOBP) terms that can be found in various intestinal sections. We summarize total number of genes corresponding to a given GO term in the

microarray gene expression (gene set size, labelled as Set Size), count of genes extracted with our analysis (observed hits, labelled as Hits) and adjusted p-values (values have been rounded-up to two decimal points; adjusted p-value < 0.1). Intestine sections: Prox = proximal; Mid = middle; Dist = distal. Response types: lm = linear; log = logarithm; exp = exponential. GOBP terms showing quadratic and cubic response types have been excluded as they can be found in the middle section only (i.e., there are no spatial trends along the small intestine axis; see Table 3.2).

Section and responses	Gene set name	Gene set term	Set size	Observed hits	Adjusted p-value
Proximal (Linear)	GO:0006631	fatty acid metabolic process	68	19	<0.01
	GO:0006629	lipid metabolic process	184	33	<0.01
	GO:0006635	fatty acid beta-oxidation	21	9	<0.01
	GO:0042632	cholesterol homeostasis	33	11	<0.01
	GO:0030301	cholesterol transport	15	7	<0.01
	GO:0033344	cholesterol efflux	17	7	<0.01
	GO:0006695	cholesterol biosynthetic process	18	7	<0.01
	GO:0006637	acyl-CoA metabolic process	19	7	<0.01
	GO:0006694	steroid biosynthetic process	42	11	<0.01
	GO:0016126	sterol biosynthetic process	16	6	<0.01
	GO:0022900	electron transport chain	69	14	<0.01
	GO:0016192	vesicle-mediated transport	134	21	0.01
	GO:0010718	positive regulation of epithelial to mesenchymal transition	15	5	0.01
	GO:0015031	protein transport	386	45	0.01
	GO:0042157	lipoprotein metabolic process	15	5	0.01
	GO:0008203	cholesterol metabolic process	50	10	0.02
	GO:0034968	histone lysine methylation	22	6	0.02
	GO:0071300	cellular response to retinoic acid	17	5	0.02
	GO:0007040	lysosome organization	18	5	0.03
	GO:0055085	transmembrane transport	390	42	0.07
	GO:0006665	sphingolipid metabolic process	15	4	0.08
	GO:0006953	acute-phase response	22	5	0.09
Proximal (Logarithm)	GO:0006917	induction of apoptosis	131	16	0.06
Middle (Linear)	GO:0006631	fatty acid metabolic process	67	28	<0.01
	GO:0006629	lipid metabolic process	183	46	<0.01
	GO:0006635	fatty acid beta-oxidation	21	11	<0.01
	GO:0005975	carbohydrate metabolic process	141	32	<0.01
	GO:0022900	electron transport chain	69	18	<0.01
	GO:0015031	protein transport	387	60	<0.01
	GO:0042542	response to hydrogen peroxide	18	7	0.01
	GO:0009058	biosynthetic process	39	11	0.01
	GO:0043154	negative regulation of caspase activity	36	10	0.02
	GO:0055085	transmembrane transport	391	57	0.02

	GO:0050796	regulation of insulin secretion	21	7	0.02
	GO:0006006	glucose metabolic process	38	10	0.03
	GO:0007040	lysosome organization	18	6	0.04
	GO:0045471	response to ethanol	29	8	0.04
	GO:0006637	acyl-CoA metabolic process	19	6	0.04
	GO:0045444	fat cell differentiation	24	7	0.04
	GO:0016042	lipid catabolic process	69	14	0.05
	GO:0032091	negative regulation of protein binding	15	5	0.05
	GO:0043161	proteasomal ubiquitin-dependent protein catabolic process	32	8	0.07
	GO:0006520	cellular amino acid metabolic process	16	5	0.07
	GO:0034612	response to tumor necrosis factor	16	5	0.07
	GO:0001890	placenta development	22	6	0.08
	GO:0006953	acute-phase response	22	6	0.08
	GO:0051262	protein tetramerization	17	5	0.09
Middle (Logarithm)	GO:0071230	cellular response to amino acid stimulus	22	7	<0.01
	GO:0007160	cell-matrix adhesion	40	8	<0.01
	GO:0051259	protein oligomerization	22	6	<0.01
	GO:0007155	cell adhesion	341	25	<0.01
	GO:0006644	phospholipid metabolic process	17	4	0.01
	GO:0018149	peptide cross-linking	17	4	0.01
	GO:0045859	regulation of protein kinase activity	16	4	0.01
	GO:0034446	substrate adhesion-dependent cell spreading	18	4	0.01
	GO:0043065	positive regulation of apoptosis	165	14	0.01
	GO:0007229	integrin-mediated signaling pathway	55	7	0.02
	GO:0006749	glutathione metabolic process	24	4	0.04
	GO:0001937	negative regulation of endothelial cell proliferation	16	3	0.06
	GO:0030574	collagen catabolic process	16	3	0.06
	GO:0009411	response to UV	29	4	0.07
	GO:0043627	response to estrogen stimulus	29	4	0.07
	GO:0045766	positive regulation of angiogenesis	57	6	0.07
	GO:0030334	regulation of cell migration	30	4	0.08
	GO:0030168	platelet activation	19	3	0.09
	GO:0007179	transforming growth factor beta receptor signaling pathway	47	5	0.10
	GO:0042632	cholesterol homeostasis	33	4	0.10
	GO:0050731	positive regulation of peptidyl-tyrosine phosphorylation	47	5	0.10
	GO:0051897	positive regulation of protein kinase B signaling cascade	33	4	0.10
Middle (Exponential)	GO:0007243	intracellular protein kinase cascade	56	8	0.03
	GO:0030301	cholesterol transport	15	4	0.03
	GO:0006468	protein phosphorylation	426	27	0.08
	GO:0007049	cell cycle	411	26	0.08
	GO:0009615	response to virus	59	7	0.08
	GO:0031532	actin cytoskeleton reorganization	22	4	0.08
	GO:0006919	activation of caspase activity	50	6	0.09
	GO:0007067	mitosis	166	13	0.09
	GO:0008630	DNA damage response. signal transduction resulting in induction of apoptosis	15	3	0.09

Distal (Linear)	GO:0051260	protein homooligomerization	70	7	<0.01
	GO:0006953	acute-phase response	21	3	0.03
	GO:0016042	lipid catabolic process	69	5	0.03
	GO:0043085	positive regulation of catalytic activity	25	3	0.03
	GO:0006099	tricarboxylic acid cycle	16	2	0.05
	GO:0006520	cellular amino acid metabolic process	16	2	0.05
	GO:0006644	phospholipid metabolic process	17	2	0.05
	GO:0007200	activation of phospholipase C activity by G-protein coupled receptor protein signaling pathway coupled to IP3 second messenger	17	2	0.05
	GO:0008654	phospholipid biosynthetic process	37	3	0.05
	GO:0033344	cholesterol efflux	17	2	0.05
	GO:0042157	lipoprotein metabolic process	16	2	0.05
	GO:0045859	regulation of protein kinase activity	16	2	0.05
	GO:0051262	protein tetramerization	17	2	0.05
	GO:0051402	neuron apoptosis	16	2	0.05
	GO:0060271	cilium morphogenesis	16	2	0.05
	GO:0006869	lipid transport	43	3	0.06
	GO:0006958	complement activation. classical pathway	20	2	0.06
	GO:0008202	steroid metabolic process	41	3	0.06
	GO:0030307	positive regulation of cell growth	44	3	0.06
	GO:0006888	ER to Golgi vesicle-mediated transport	22	2	0.06
	GO:0009725	response to hormone stimulus	22	2	0.06
	GO:0006749	glutathione metabolic process	24	2	0.08
	GO:0016049	cell growth	24	2	0.08
	GO:0007224	smoothened signaling pathway	25	2	0.08
	GO:0008203	cholesterol metabolic process	50	3	0.08
	GO:0045785	positive regulation of cell adhesion	25	2	0.08
Distal (Exponential)	GO:0000910	cytokinesis	31	4	<0.01
	GO:0001824	blastocyst development	16	2	0.01
	GO:0043967	histone H4 acetylation	15	2	0.01
	GO:0019882	antigen processing and presentation	18	2	0.02
	GO:0030218	erythrocyte differentiation	24	2	0.03
	GO:0032313	regulation of Rab GTPase activity	27	2	0.03
	GO:0032851	positive regulation of Rab GTPase activity	28	2	0.03
	GO:0043154	negative regulation of caspase activity	36	2	0.06

Table 3.5: Over-represented Gene Ontology Biological Process (GOBP) terms associated to differentially expressed genes in the three intestinal sections. Genes are classified as linear, logarithmic or exponential, according to the best (i.e., with the smallest adjusted p-value) response type describing their expression pattern. For each significant GOBP term (described by Gene set name and Gene set term) we summarize number of genes corresponding to the process (Set size), number of genes found with our analysis (Observed hits) and adjusted p-value. Universe size: proximal = 14,952; middle = 14,933; distal = 14,925.

Sections and responses	GOBP	Proximal significant genes	Middle significant genes	Distal significant genes
Prox – Mid – Dist (all linear)	GO:0006953: acute-phase response	UP: *Saa3*, *Serpina1b* DOWN: *Reg3b*, *Reg3g*, *Stat3*	UP: *Serpina1b* DOWN: *Reg3b*, *Reg3g*, *Saa2*, *Stat3*, *Sigirr*	UP: *Serpina1b* DOWN: *Saa1*, *Saa2*
Prox – Mid (all linear)	GO:0006629: lipid metabolic process	UP: *Acadl*, *Acadvl*, *Acads*, *Apoc2*, *Cpt2*, *Crat*, *Ldlr*, *Acer1*, *Acot2*, *Acot4*, *Acaa1b*, *Slc27a2*, *Acot1*, *Plcxd1*, *Ech1*, *Acaa2*, *Gde1*, *Hacl1*, *Pnpla8*, *Abhd5*, *Adipor2*, *Arv1*, *Crot*, *Acsl3*, *Echs1*, *Hadha* DOWN: *Plcb3*, *Daglb*, *Scap*, *Npc1l1*, *Apob*, *Asah2*, *Serinc1*	UP: *Tecr*, *Hdlbp*, *Acadl*, *Acadvl*, *Acox1*, *Cpt1a*, *Cpt2*, *Crat*, *Fdxr*, *Gpx1*, *Hsd17b4*, *Lipa*, *Acer1*, *Acot4*, *Mttp*, *Pck1*, *Soat2*, *Acaa1b*, *Slc27a2*, *Slc27a4*, *Acot1*, *Hsd17b6*, *Acsl5*, *Ech1*, *Echdc2*, *Acaa2*, *Gde1*, *Hacl1*, *Lpin2*, *Lpin3*, *Pnpla8*, *Adipor2*, *Acot12*, *Gpcpd1*, *Acsl3*, *Acox2*, *Echs1*, *Hadha* DOWN: *Pafah1b2*, *Ppard*, *Daglb*, *Gdpd5*, *Napepld*, *Asah2*, *Plb1*, *Plce1*	
	GO:0006631: fatty acid metabolic process	UP: *Acadl*, *Acadvl*, *Acads*, *Apoa2*, *Cd36*, *Cpt2*, *Crat*, *Prkar2b*, *Acaa1b*, *Slc27a2*, *Ech1*, *Acaa2*, *Pnpla8*, *Abhd5*, *Adipor2*, *Crot*, *Acsl3*, *Echs1*, *Hadha*	UP: *Acadl*, *Acadvl*, *Acox1*, *Cd36*, *Cpt1a*, *Cpt2*, *Crat*, *Faah*, *Fabp2*, *Hsd17b4*, *Lipa*, *Acaa1b*, *Slc27a2*, *Slc27a4*, *Mecr*, *Acsl5*, *Ech1*, *Echdc2*, *Acaa2*, *Lpin2*, *Lpin3*, *Pnpla8*, *Adipor2*, *Acot12*, *Acsl3*, *Acox2*, *Echs1*, *Hadha*	
	GO:0006635: fatty acid beta-oxidation	UP: *Acadvl*, *Eci1*, *Slc25a17*, *Eci2*, *Acaa2*, *Decr1*, *Bdh2*, *Echs1*, *Hadha*	UP: *Acadvl*, *Acox1*, *Hsd17b4*, *Pex5*, *Slc25a17*, *Eci2*, *Acaa2*, *Bdh2*, *Acox2*, *Echs1*, *Hadha*	
	GO:0006637: acyl-CoA metabolic process	UP: *Hmgcl*, *Acot8*, *Acot2*, *Acot4*, *Acot1*, *Oxsm* DOWN: *Ces1d*	UP: *Hmgcl*, *Acot8*, *Acot4*, *Acot1*, *Oxsm*, *Acot12* DOWN:	
	GO:0007040: lysosome organization	UP: *Cln8* DOWN: *Abca1*, *Tpp1*, *Hexa*, *Cln6*	UP: *Acp2*, *Fam160a2* DOWN: *Tpp1*, *Hexa*, *Hexb*, *Ppt1*	
	GO:0015031: protein transport	UP: *Kdelr3*, *Bet1*, *Nacad*, *Rab1*, *Rab6*, *Mcfd2*, *Sec22b*, *Sec23a*, *Timm17b*, *Vps45*, *Atg4d*, *Rabep1*, *Rab9*, *Vps29*, *Apba3*, *Cope*, *Xpo7*, *Sar1b*, *Gpr89*, *Yif1a*, *Golt1a*, *Snx4*, *Pex13*, *Senp2*, *Snx7* DOWN: *Sft2d2*, *Gdi1*, *Cog1*, *Myo1c*, *Rab12*, *Vps33b*, *Rab8b*, *Scamp2*, *Gga3*, *Agap1*, *Copg2*, *Nup210*, *Sec61a2*, *Rab37*, *Rab32*, *Kdelr1*, *Ap2b1*, *Gga2*, *Chmp4b*, *Rrbp1*	UP: *Chmp7*, *Kdelr3*, *Sec13*, *Vps4a*, *Ap1m1*, *Arf4*, *Bet1*, *Pex5*, *Sec22b*, *Sec23a*, *Stam*, *Arcn1*, *Timm17a*, *Timm17b*, *Cox18*, *Bcap31*, *Eif5a*, *Serp1*, *Timm13*, *Ipo8*, *Uevld*, *Copg*, *Pex14*, *Copz1*, *Gosr2*, *Apba3*, *Rab2a*, *Cope*, *Tomm6*, *Sar1b*, *Kdelr2*, *Derl1*, *Yif1a*, *Golt1a*, *Sec62*, *Dopey2*, *Tmem48*, *Fam160a2*, *Senp2*, *Sil1* DOWN: *Arrb1*, *Ap3d1*, *Gdi1*, *Mtm1*, *Ppt1*, *Rab17*, *Rab19*, *Rab33b*, *Rab3ip*, *Rufy1*, *Vps33b*, *Rab8b*, *Sec61a2*, *Rab37*, *Tomm40l*, *Duoxa2*, *Rab32*, *Rab27b*, *Rrbp1*, *Cog8*	
	GO:0022900: electron transport chain	UP: *Ndufb11*, *Etfb*, *Etfa*, *Ndufa4*, *Ndufs8*, *Ndufs6*, *Cyb561d2*, *Ndufa7*, *Cyb5b*, *Uqcrfs1*, *Etfdh*	UP: *Etfb*, *Etfa*, *Fdxr*, *Ndufs8*, *Ndufs2*, *Ndufb6*, *Ndufs6*, *Cyb561d2*, *Ndufs5*, *Ndufa9*, *Cyb5b*,	

		DOWN: *Cmah, Cyb561, Cyba*	*Uqcrfs1, Etfdh, Ndufa5* DOWN: *Cyb561, Frrs1, Enox2, Cyb561d1*	
	GO:0055085: transmembrane transport	UP: *Slc30a9, Aqp7, Aqp8, Cacna1d, Slc6a3, Sfxn1, Slc22a1, Slc23a1, Slc25a17, Slc6a2, Mfsd6l, Timm17b, Slc6a20b, Slc4a11, Gpr172b, Xpo7, Slc16a13, Abcb6, Senp2, Mfsd2a* DOWN: *Abca1, Atp8a1, Slc7a1, Itpr3, Myo1c, Slc12a7, Slc19a1, Slc2a2, Mfsd4, Slc36a1, Slc7a9, Atp2a3, Slc4a4, Nup210, Abcb10, Sec61a2, Cacna1h, Slc30a7, Slc39a8, Slc29a3, Hvcn1, Rrbp1*	UP: *Slc25a45, Slc30a9, Slc35b1, Sec13, Aqp3, Aqp7, Abcc2, Slc6a3, Slc26a3, Sfxn1, Slc1a3, Slc22a1, Slc23a1, Slc25a17, Slc7a7, Mfsd7c, Timm17a, Timm17b, Catsper1, Slc6a20b, Slc16a12, Mfsd7a, Eif5a, Serp1, Timm13, Gpr172b, Pex14, Slc47a1, Sec62, Slc16a13, Slc39a11, Slc39a5, Tmem48, Abcb6, Abcb8, Senp2, Sil1* DOWN: *Slc26a2, Kcnu1, Slc24a6, Slc1a1, Slc2a2, Slc5a1, Spns2, Trpm6, Slc5a9, Slc17a5, Slc7a9, Atp9b, Sec61a2, Tomm40l, Hiatl1, Mfsd1, Spns1, Hvcn1, Rrbp1, Slc4a10*	
Prox – Mid (linear - logarithmic)	GO:0042632: cholesterol homeostasis	UP: *Apoa2, Ldlr, Nr1h3, Pla2g10, Npc2* DOWN: *Abca1, Npc1, Npc1l1, Apob, Abcg5, Abcg8*	UP: *Apoa4, Cav1, Pla2g10* DOWN: *Cyp7a1*	
Prox – Mid (linear - exponential)	GO:0030301: cholesterol transport	UP: *Apoa2, Cd36, Ldlr, Npc2* DOWN: *Abca1, Npc1l1, Apob*	UP: DOWN: *Abca1, Abcg1, Scarb1, Stard3*	
Mid – Dist (all linear)	GO:0006520: cellular amino acid metabolic process		UP: *Cth, Acy1, Ddc, Got1* DOWN: *Ccbl2*	UP: *Cth* DOWN: *Tat*
	GO:0016042: lipid catabolic process		UP: *Lipa, Pla2g2c, Nceh1, Pnpla8, Ddhd2* DOWN: *Ddhd1, Pafah1b2, Pla2g2a, Pld2, Ppt1, Daglb, Plb1, Plbd1, Plce1*	UP: Apoc2 DOWN: Pla2g2a, Daglb, Pla2g2f, Plce1
	GO:0051262: protein tetramerization		UP: *Me1, Pex5, Txnrd1* DOWN: *Igf1r, Sbf2*	UP: *Me1* DOWN: *Sbf2*
Mid – Dist (linear - exponential)	GO:0043154: negative regulation of caspase activity		UP: *Prdx3, Birc5, Gpx1, Igf1, Por* DOWN: *Arrb1, Bcl2l1, Naip1, Igbp1, Usp47*	UP: *Birc5* DOWN: *Igbp1*
Mid – Dist (logarithmic - linear)	GO:0006644: phospholipid metabolic process		UP: *Pla2g2d, Pla2g10, Pla2g12a, Ppap2b*	DOWN: *Pla2g2a, Pla2g2f*
	GO:0006749: glutathione metabolic process		UP: *Gpx3, Gstm1, Gstm3, Gstk1* DOWN:	UP: *Cth, Hagh* DOWN:
	GO:0045859: regulation of protein kinase activity		UP: *Prkar2b, Mtor* DOWN: *Tsc1, Akap13*	UP: *Ect2* DOWN: *Plce1*

Table 3.6: Over-represented Gene Ontology Biological Process (GOBP) terms that can be found in various intestinal sections. For each GOBP term we specify up- and down-regulated genes (adjusted p-value < 0.1; most of the results of gene set enrichment analysis fall well below this threshold - see Table 3.5).

Section and responses	GO ID	GOBP	Proximal			Middle			Distal		
			Set size	Hits	Adjusted p-value	Set size	Hits	Adjusted p-value	Set size	Hits	Adjusted p-value
Prox – Mid – Dist (all linear)	GO:0006749	glutathione metabolic process	24	5	0.01	24	6	0.01	24	2	0.04
	GO:0006888	ER to Golgi vesicle-mediated transport	22	4	0.05	22	5	0.02	22	2	0.03
Prox – Mid – Dist (lm – log - lm)	GO:0006749	glutathione metabolic process	24	5	0.01	24	4	<0.01	24	2	0.04
	GO:0042632	cholesterol homeostasis	33	5	0.06	33	3	0.04	33	2	0.08
Prox – Mid (lm - lm)	GO:0006629	lipid metabolic process	184	26	<0.01	183	38	<0.01			
	GO:0006631	fatty acid metabolic process	68	19	<0.01	67	28	<0.01			
	GO:0006635	fatty acid beta-oxidation	21	9	<0.01	21	11	<0.01			
	GO:0006637	acyl-CoA metabolic process	19	6	<0.01	19	6	<0.01			
	GO:0006695	cholesterol biosynthetic process	18	5	<0.01	18	4	0.04			
	GO:0008610	lipid biosynthetic process	81	10	0.01	81	11	0.02			
	GO:0016192	vesicle-mediated transport	134	13	0.03	134	19	<0.01			
	GO:0022900	electron transport chain	69	11	<0.01	69	14	<0.01			
	GO:0045454	cell redox homeostasis	47	6	0.07	47	9	<0.01			
	GO:0051289	protein homotetramerization	38	7	<0.01	38	6	0.07			
Prox – Mid (lm - log)	GO:0022900	electron transport chain	69	11	<0.01	69	5	0.03			
Prox – Mid (log - lm)	GO:0006631	fatty acid metabolic process	68	6	0.07	67	28	<0.01			
Mid – Dist (log - lm)	GO:0033344	cholesterol efflux				17	2	0.04	17	2	0.02
	GO:0042325	regulation of phosphorylation				34	3	0.04	34	2	0.08

Table 3.7: Over-represented Gene Ontology Biological Process (GOBP) terms completely up-regulated and found in various intestinal sections. For each GOBP term we summarize: the total number of genes in the microarray (gene set size, labelled as Set size), the count of genes extracted with our analysis (observed hits, labelled as Hits), and the adjusted p-values. Intestine sections: Prox = proximal; Mid = middle; Dist = distal. Response types: lm = linear; log = logarithm; exp = exponential.

Section and responses	GOBP	Proximal significant genes	Middle significant genes	Distal significant genes
Prox – Mid – Dist (all linear)	GO:0006749: glutathione metabolic process	*Cth, Glo1, Hagh, Gstz1, Idh1*	*Cth, G6pdx, Ggt1, Hagh, Gpx1, Gstz1*	*Cth, Hagh*
	GO:0006888: ER to Golgi vesicle-mediated transport	*Bet1, Rab1, Sec22b, Sec23a*	*Sec13, Bet1, Sec22b, Sec23a, Lmf1*	*Sec22b, Lman1*
Prox – Mid – Dist (linear – logarithmic - linear)	GO:0006749: glutathione metabolic process	*Cth, Glo1, Hagh, Gstz1, Idh1*	*Gpx3, Gstm1, Gstm3, Gstk1*	*Cth, Hagh*
	GO:0042632: cholesterol homeostasis	*Apoa2, Ldlr, Nr1h3, Pla2g10, Npc2*	*Apoa4, Cav1, Pla2g10*	*Apoa4, Mttp*
Prox – Mid (all linear)	GO:0006629: lipid metabolic process	*Acadl, Acadvl, Acads, Apoc2, Cpt2, Crat, Ldlr, Acer1, Acot2, Acot4, Acaa1b, Slc27a2, Acot1, Plcxd1, Ech1, Acaa2, Gde1, Hacl1, Pnpla8, Abhd5, Adipor2, Arv1, Crot, Acsl3, Echs1, Hadha*	*Tecr, Hdlbp, Acadl, Acadvl, Acox1, Cpt1a, Cpt2, Crat, Fdxr, Gpx1, Hsd17b4, Lipa, Acer1, Acot4, Mttp, Pck1, Soat2, Acaa1b, Slc27a2, Slc27a4, Acot1, Hsd17b6, Acsl5, Ech1, Echdc2, Acaa2, Gde1, Hacl1, Lpin2, Lpin3, Pnpla8, Adipor2, Acot12, Gpcpd1, Acsl3, Acox2, Echs1, Hadha*	
	GO:0006631: fatty acid metabolic process	*Acadl, Acadvl, Acads, Apoa2, Cd36, Cpt2, Crat, Prkar2b, Acaa1b, Slc27a2, Ech1, Acaa2, Pnpla8, Abhd5, Adipor2, Crot, Acsl3, Echs1, Hadha*	*Acadl, Acadvl, Acox1, Cd36, Cpt1a, Cpt2, Crat, Faah, Fabp2, Hsd17b4, Lipa, Acaa1b, Slc27a2, Slc27a4, Mecr, Acsl5, Ech1, Echdc2, Acaa2, Lpin2, Lpin3, Pnpla8, Adipor2, Acot12, Acsl3, Acox2, Echs1, Hadha*	
	GO:0006635: fatty acid beta-oxidation	*Acadvl, Eci1, Slc25a17, Eci2, Acaa2, Decr1, Bdh2, Echs1, Hadha*	*Acadvl, Acox1, Hsd17b4, Pex5, Slc25a17, Eci2, Acaa2, Bdh2, Acox2, Echs1, Hadha*	
	GO:0006637: acyl-CoA metabolic process	*Hmgcl, Acot8, Acot2, Acot4, Acot1, Oxsm*	*Hmgcl, Acot8, Acot4, Acot1, Oxsm, Acot12*	
	GO:0006695: cholesterol biosynthetic process	*Ebp, Hmgcs2, Nsdhl, Pmvk, Tm7sf2*	*Ebp, G6pdx, Hmgcs2, Tm7sf2*	
	GO:0008610: lipid biosynthetic process	*Hsd17b11, Ebp, H2-Ke6, Hmgcs2, Hsd17b2, Nsdhl, Acss2, Pmvk, Oxsm, Tm7sf2*	*Tecr, Hsd17b11, Ebp, Hmgcs2, Mecr, Hsd17b12, Acss2, Ptplad2, Oxsm, Tm7sf2, Lass2*	
	GO:0016192: vesicle-mediated transport	*Kdelr3, Bet1, Rab1, Rab6, Mcfd2, Sec22b, Sec23a, Vps45, Cope, Sar1b, Ap4b1, Yif1a, Golt1a*	*Kdelr3, Sec13, Vps4a, Ap1m1, Arf4, Bet1, Sec22b, Sec23a, Arcn1, Bcap31, Copg, Copz1, Gosr2, Rab2a, Cope, Sar1b, Kdelr2, Yif1a, Golt1a*	
	GO:0022900: electron transport chain	*Ndufb11, Etfb, Etfa, Ndufa4, Ndufs8, Ndufs6, Cyb561d2, Ndufa7, Cyb5b, Uqcrfs1, Etfdh*	*Etfb, Etfa, Fdxr, Ndufs8, Ndufs2, Ndufb6, Ndufs6, Cyb561d2, Ndufs5, Ndufa9, Cyb5b, Uqcrfs1, Etfdh, Ndufa5*	
	GO:0045454: cell redox homeostasis	*1810046J19Rik, Ddit3, Txnrd2, 2810407C02Rik, Txndc15, Glrx5*	*1810046J19Rik, Gpx1, Txnrd1, 2810407C02Rik, Txndc15, Pdia6, Pdia5, Glrx5, Erp44*	
	GO:0051289: protein homotetramerization	*Acacb, Cth, Acadl, Acads, Aldoc, Dhps, Decr1*	*Cth, Acadl, Aldh1a1, Aldoc, Ide, Cda*	
Prox – Mid (linear - logarithmic)	GO:0022900: electron transport chain	*Ndufb11, Etfb, Etfa, Ndufa4, Ndufs8, Ndufs6, Cyb561d2, Ndufa7, Cyb5b, Uqcrfs1, Etfdh*	*Txn2, Ndufb9, Ndufc1, Ndufa7, Steap2*	

Prox- Mid (logarithmic - linear)	GO:0006631: fatty acid metabolic process	*Fabp2, Hsd17b4, Slc27a4, Acsl5, Acot12, Acox2*	*Acadl, Acadvl, Acox1, Cd36, Cpt1a, Cpt2, Crat, Faah, Fabp2, Hsd17b4, Lipa, Acaa1b, Slc27a2, Slc27a4, Mecr, Acsl5, Ech1, Echdc2, Acaa2, Lpin2, Lpin3, Pnpla8, Adipor2, Acot12, Acsl3, Acox2, Echs1, Hadha*	
Mid – Dist (logarithmic - linear)	GO:0033344: cholesterol efflux		*Apoa4, Cav1*	*Apoa4, Apoc2*
	GO:0042325: regulation of phosphorylation		*Pik3r3, Srpx2, Rptor*	*Pik3r3, Rptor*

Table 3.8: Over-represented Gene Ontology Biological Process (GOBP) terms that include up-regulated genes only. Such GOBP terms can be found in various intestinal sections; lists of (up-regulated) differentially expressed genes associated to each GOBP term are summarized (adjusted p-value < 0.1).

Section and responses	GO ID	GOBP	Proximal			Middle			Distal		
			Set size	Hits	Adjusted p-value	Set size	Hits	Adjusted p-value	Set size	Hits	Adjusted p-value
Prox – Mid – Dist (all linear)	GO:0045785	positive regulation of cell adhesion	26	4	0.05	26	5	0.08	25	2	0.03
Prox – Mid – Dist (lm – exp - lm)	GO:0003333	amino acid transmembrane transport	39	5	0.05	39	3	0.09	39	2	0.07
Prox – Mid (lm - lm)	GO:0007040	lysosome organization	18	4	0.02	18	4	0.08			
	GO:0009615	response to virus	59	7	0.03	59	9	0.07			
	GO:0030335	positive regulation of cell migration	83	10	0.01	83	10	0.08			
Prox – Mid (lm - exp)	GO:0007243	intracellular protein kinase cascade	56	7	0.03	56	6	0.01			
	GO:0009615	response to virus	59	7	0.03	59	5	0.04			
	GO:0030301	cholesterol transport	15	3	0.05	15	4	<0.01			
	GO:0033344	cholesterol efflux	17	4	0.01	17	3	0.03			
	GO:0042632	cholesterol homeostasis	33	6	0.01	33	3	0.07			
	GO:0046777	protein autophosphorylation	91	9	0.04	90	5	0.09			
	GO:0071300	cellular response to retinoic acid	17	5	<0.01	17	3	0.03			
Prox – Mid (log - lm)	GO:0009615	response to virus	59	6	0.04	59	9	0.07			
Prox – Mid (log - log)	GO:0043029	T cell homeostasis	15	3	0.03	15	2	0.08			
Prox – Mid (log - exp)	GO:0009615	response to virus	59	6	0.04	59	5	0.04			
	GO:0016358	dendrite development	18	4	0.01	20	2	0.09			
Mid - Dist (lm - lm)	GO:0006953	acute-phase response				22	5	0.07	21	2	0.02
	GO:0016042	lipid catabolic process				69	9	0.08	69	4	0.02

Table 3.9: Over-represented Gene Ontology Biological Process (GOBP) terms with down-regulated genes only. Data refer to GOBP terms that can be found in various intestinal sections. For each GOBP term

we indicate: total number of genes in the microarray (gene set size, labelled as Set size), count of genes extracted with our analysis (observed hits, labelled as Hits), and adjusted p-values. Intestine sections: Prox = proximal; Mid = middle; Dist = distal. Response types: lm = linear; log = logarithm; exp = exponential.

Section and responses	GOBP	Proximal significant genes	Middle significant genes	Distal significant genes
Prox – Mid – Dist (all linear)	GO:0045785: positive regulation of cell adhesion	*Fgf1, Cx3cl1, Tgm2, Vegfa*	*Fgf1, Itgav, Pld2, Cx3cl1, Tgm2*	*Saa1, Tgm2*
Prox – Mid – Dist (linear – exponential - linear)	GO:0003333: amino acid transmembrane transport	*Slc7a1, Slc1a5, Slc36a1, Slc7a9, Serinc1*	*Slco4a1, Slc3a2, Slc36a1*	*Slc1a1, Slc7a9*
Prox – Mid (all linear)	GO:0007040: lysosome organization	*Abca1, Tpp1, Hexa, Cln6*	*Tpp1, Hexa, Hexb, Ppt1*	
	GO:0009615: response to virus	*Pml, Mst1r, Mavs, Ddx58, Oas1b, Unc93b1, Ifih1*	*Bcl2l1, Ifit1, Ddx58, Rnasel, Trim56, Irf7, Zbp1, Ifih1, Zc3hav1*	
	GO:0030335: positive regulation of cell migration	*Acp5, Adam10, Adam17, Bmp2, Fgf1, Igf1r, Ptk2b, Cx3cl1, Vegfa, Zfp703*	*Adam10, Bmp2, Fgf1, Igf1r, Itgav, Pik3r1, Pld2, Ptk2b, Cx3cl1, Zfp703*	
Prox – Mid (linear - exponential)	GO:0007243: intracellular protein kinase cascade	*Socs3, Fgf1, Smpd2, Wnk1, Rps6ka4, Tnik, Dapk1*	*Socs3, Jak2, Mknk1, Slc9a1, Ick, Tnik*	
	GO:0009615: response to virus	*Pml, Mst1r, Mavs, Ddx58, Oas1b, Unc93b1, Ifih1*	*Mx1, Mx2, Pml, Mst1r, Mavs*	
	GO:0030301: cholesterol transport	*Abca1, Npc1l1, Apob*	*Abca1, Abcg1, Scarb1, Stard3*	
	GO:0033344: cholesterol efflux	*Abca1, Npc1, Abcg5, Abcg8*	*Abca1, Abcg1, Scarb1*	
	GO:0042632: cholesterol homeostasis	*Abca1, Npc1, Npc1l1, Apob, Abcg5, Abcg8*	*Abca1, Abcg1, Scarb1*	
	GO:0046777: protein autophosphorylation	*Clk2, Igf1r, Uhmk1, Lyn, Pak2, Wnk1, Map3k1, Taok3, Tnik*	*Jak2, Lyn, Pak1, Pim3, Tnik*	
	GO:0071300: cellular response to retinoic acid	*Abca1, Lyn, Ptk2b, Ptk6, Mll5*	*Abca1, Lyn, Rxrb*	
Prox –Mid (logarithmic - linear)	GO:0009615: response to virus	*Ifit1, Rnasel, Trim56, Irf7, Zbp1, Polr3f*	*Bcl2l1, Ifit1, Ddx58, Rnasel, Trim56, Irf7, Zbp1, Ifih1, Zc3hav1*	
Prox – Mid (all logarithmic)	GO:0043029: T cell homeostasis	*Bcl2l11, Fas, Stat5b*	*Bcl2l11, Gimap5*	
Prox – Mid (logarithmic - exponential)	GO:0009615: response to virus	*Ifit1, Rnasel, Trim56, Irf7, Zbp1, Polr3f*	*Mx1, Mx2, Pml, Mst1r, Mavs*	
	GO:0016358: dendrite development	*Mecp2, Pak1, Reln, Abi2*	*Pak1, Ss18l1*	
Mid – Dist (all linear)	GO:0006953: acute-phase response		*Reg3b, Reg3g, Saa2, Stat3, Sigirr*	*Saa1, Saa2*
	GO:0016042: lipid catabolic process		*Ddhd1, Pafah1b2, Pla2g2a, Pld2, Ppt1, Daglb, Plb1, Plbd1, Plce1*	*Pla2g2a, Daglb, Pla2g2f, Plce1*

Table 3.10: Over-represented down-regulated Gene Ontology Biological Process (GOBP) terms that can be found in various intestinal sections. For each term we specify the associated differentially expressed genes (adjusted p-value < 0.1).

3.4 Discussion

Excessive saturated fat intake leads to metabolic disorders (Balwierz et al. 2009; Cheung & Mao 2012; Desmarchelier et al. 2012; Walley et al. 2006; Zhang et al. 2009). Especially in the small intestine, studies have shown connections to altered lipid metabolic functions or illustrated its role in the development of obesity and diabetes (Kondo et al. 2006; de Wit et al. 2008). However, it is quite rare to find analyses on intestinal section-specific response to dietary fat. Differences in the absorption capabilities of intestine sections are caused by changes in gene expression (Chen et al. 2001; Simon et al. 2011). de Wit and colleagues analyzed the transcriptomic variation between dietary groups (e.g., 20% vs. 10%, 45% vs. 10% calories from fat), specific to three sections of small intestine of mice (de Wit et al. 2011). In their experiment, mice were fed diets that differed in dietary fat and carbohydrate content; the respective diets contained 10%, 20%, 30% or 45% kcal fat, which was termed as fat intake (within the field of nutrition it is generally accepted to refer to macronutrient content of the diets as percentage of calories derived from fat). Our goal was to re-analyze their data to assess linear and nonlinear transcriptomic response to fat intake by modeling gene expression as a function of fat intake %. The percentage of calories derived from fat intake represented the independent variable, even though such percentage might have been different from the relative amount of calories from fat that was effectively sensed by enterocytes. Moreover, the absorption capacity of enterocytes could be modulated by dietary fat (but in our study the type of fat that varied was always palm oil; i.e., C16:0).

3.4.1 Gene- and section-specific transcriptomic response to dietary fat intake

We investigated whether the transcriptomic response type was related to specific intestine sections (thus being potentially triggered by fat availability along the longitudinal axis), or uniquely associated to specific genes and biological processes (and preserved in all

intestine sections, without any change along the longitudinal axis). Due to progressive absorption along the intestine, we expected higher amount of fat to be available in the proximal section and a progressive decrease towards the distal section. Furthermore, it can be the case that the quantity of dietary fat exceeds the absorption capacity of the proximal section, and thus "overflows" to the middle and distal sections. Moreover, the middle section has greater absorptive area than proximal and distal sections, and is known to be highly affected by the dietary fat (de Wit et al. 2008). Our results corroborate previous findings, confirming that the highest number of differentially expressed genes is in the middle section of small intestine (the proximal part has a relevant metabolic role, while the middle and distal sections are mainly dedicated to fat absorption and transportation processes). However, two new and clear patterns emerged from our analysis: (I) most of the genes characterized by a significant relationship between expression levels and dietary fat intake exhibited linear responses (i.e., their activity was not overloaded, even at high fat concentration), and the prevalence of linear relationships was persistent along the whole small intestine; (II) the relevance of nonlinear relationships linking fat intake to gene expression levels was section-specific and reflected the progressive reduction of fat availability (due to intestinal absorption) along the longitudinal axis of the small intestine (i.e., the relative importance of logarithmic responses, which can be associated with overload mechanisms, decreased from proximal to distal section, while the relative importance of exponential responses increased when moving towards the distal part). The relevance of quadratic and cubic response types was marginal both in absolute and relative terms, in all sections. Throughout the small-intestine, we observed biological processes that responded to fat intake in a single section, and other processes with either unaltered (e.g., acute-phase response is linear in all three sections; Table 3.4) or varying responses across the sections (e.g., cholesterol transport changes from linear response in the proximal

section, with both up- and down-regulated genes, to exponential response in the middle section, with down-regulated genes only; Table 3.6).

(1) Gene-specific acute-phase response to fat. When analyzing differentially expressed genes, we observed that over-representation of APR was linked to fat intake through a linear relationship in all three intestinal sections (Table 3.4). Increasing dietary fat intake caused intestinal inflammatory transcriptomic response and led to linear increase in APR. This response does not show a univocal pattern and is associated to both up- and down-regulated genes (e.g., *Serpina1b* is up-regulated; *Reg3b* and *Reg3g* are down-regulated). Serine protease inhibitor (*Serpina1b*) was found to be differentially expressed and up-regulated in all three sections. Some studies have shown that the presence of lipopolysaccharide (LPS) in the small intestine stimulates pro-inflammatory mediators that are activators of insulin resistance (Cani et al. 2007; Ding et al. 2010) (Cani et al. 2007; Ding et al. 2010). The resultant inflammatory reaction causes APR and its prolonged activation is seen with increased plasma levels of small, low-density lipoproteins (LDLs; Khovidhunkit et al. 2000, 2003). Moreover, during inflammation plasma triglyceride and very low-density lipoprotein (VLDL) levels rise, while high-density lipoprotein (HDL) level declines (Khovidhunkit et al. 2000). In the present study, *Reg3g* and *Reg3b* are down-regulated both in proximal and middle small intestine. The protein encoded by these genes (REG3G and REG3B) belong to the family of C-type lectins and are secreted by epithelial cells and Paneth cells. Their expression is reduced in the small intestine of mice fed alcohol compared with control mice, and the effect is more pronounced in the proximal part (Hartmann et al. 2013). REG3G and REG3B play a protective role against gram-positive and gram-negative bacteria, respectively (Abreu 2010; van Ampting et al. 2012). REG3G is essential for maintaining a zone that physically separates the luminal bacteria from the small intestinal epithelial surface (Schnabl 2013). Dysregulation of a microorganism-induced program of epithelial cell homeostasis and repair can result in

chronic inflammatory responses, and is associated with the development of colon cancer (Abreu 2010; Khovidhunkit et al. 2000). Down-regulation of *Reg3g* and *Reg3b* is likely to result in increased bacterial colonization of the intestinal epithelial surface, with consequent induction of inflammation. The oncogenic transcription factor *Stat3* is down-regulated in both the proximal and the middle section, and the same pattern holds in middle and distal sections for *Saa2*, a member of the family of APR proteins that is usually up-regulated during infection, tissue damage or inflammation disease (Ivanov et al. 2009).

(2) Section-specific cholesterol processing through the intestinal length. A complex scenario characterizes cholesterol, with a prevalence of up- and down-regulation for what concerns cholesterol homeostasis and transport, respectively (Tables 3.5 – 3.10). In the middle section, cholesterol homeostasis was overloaded (i.e., logarithmic response), while cholesterol transport was particularly impaired in presence of highest levels of fat intake (i.e., exponential response). In the middle section, we observed the up-regulated, logarithmic response in the expression of cholesterol homeostasis genes (i.e., *Apoa4*, *Cav1* and *Pla2g10*), except for what concerns *Cyp7a1* (down-regulated, logarithmic). In different parts of the small intestine, we found down-regulation of ABC transporters: *Abca1* in proximal as well as in the middle section, *Abcg1* in middle section, and *Abcg5* and *Abcg8* in the proximal section (Table 3.6).These genes displayed fat-responsive expression pattern and are involved in cholesterol transport and cholesterol efflux. They are characterized by a change in the response pattern from proximal (linear response) to middle (exponential response) section. These data suggest that high fat concentrations available in the proximal small intestine down-regulate ABC transporters involved in cholesterol efflux (*Abca1*, *Abcg5* and *Abcg8*; Tables 3.6, 3.8 and 3.10). The linear response of these down-regulated genes stands for relevant effects also in presence of low fat intakes. The negative exponential response observed in the middle section for *Abca1* and *Abcg1* illustrates how only the highest level of fat intake results in the sharp impairment of

cholesterol absorption. This is in line with the fact that the middle section is characterized by a higher uptake capacity than the proximal section. Sections that are downstream the proximal small intestine (which consists of the duodenum plus the proximal jejunum) are exposed to smaller amounts of fat and less pronounced effects (except for the case of highest fat content in the diet; i.e., 45 E%). The highest fat concentration is the only one that triggered a strong change, leading to the down-regulation of two ABC transporters in the middle section.

(3) Section-specific negative regulation of caspase activity. The negative regulation of caspase activity shows a transition from linear response in the middle section to exponential response in the distal part. This process does not present a univocal response pattern and includes both up- and down-regulated genes. In the middle and distal sections, there was up-regulation of the gene *Birc5* (baculoviral IAP repeat-containing 5, also known as survivin). *Birc5* is bifunctional, plays a key role in inhibition of apoptosis and regulation of mitosis, and is essential for cell division (Blanc-Brude et al. 2007; Chiou et al. 2003; Fukuda & Pelus 2006). *Birc5* is mostly linked with carcinogenesis; during early atherogenesis, it shows elevated expression in inflamed macrophage-rich areas (Blanc-Brude et al. 2007). Linear and positive transcriptomic response of genes *Prdx3* (that encodes the protein thioredoxin-dependent peroxide reductase, mitochondrial) and *Gpx1* (that encodes the enzyme glutathione peroxidase 1) was confined to the middle, small intestine. Bellafante et al. (2014) have shown that PGC-1β overexpression in enterocytes enhances antioxidant enzymes such as *Sod2*, *Gpx4*, *Prdx3*, *Prdx5*, *Txn2* and *Sirt3*. Up-regulation of these enzymes has an antiapoptotic role both in normal mucosa and protumorigenic conditions, and causes a greater increase in the length of the villi of the small intestine. Chu et al. (2004) have observed that the targeted disruption of *Gpx1* and *Gpx2* harms two glutathione peroxidase (GPX) isoenzymes. GPX isoenzymes reduce hydroperoxidases in intestinal epithelium and their impairment increases the sensitivity of

ileum and colon to bacteria-associated inflammation and cancer. Our results show linear, up-regulation of *Prdx3* and *Gpx1* in the middle section of small intestine; this can activate antioxidant processes that are relevant for bacteria-associated inflammation, cancer-promoting conditions and tumor progression. Our result suggests that these biological processes are not overloaded in the middle and distal sections, possibly due to the absorption that contributes, in a progressive way, to reduce the amount of fat in the small intestine lumen. The linear, positive responses found in the middle part illustrate the relevance of antiapoptotic processes, even in presence of low fat intake. The exponential, positive response identified in the distal section (specific to *Birc5*) is descriptive of an outstanding activation, only for most extreme values of fat intake.

(4) Transcriptomic response in small intestine: a comparative analysis with other metabolic organs. Our study aimed at identifying changes in differentially expressed genes in three small intestine sections as a response to dietary fat intake. Kondo et al. (2006) analyzed differential gene expression among multiple metabolic organs, including small intestine. They showed that some lipid metabolism-related genes (i.e., *Mod1*, *Cyp4a10*, *Hmgcs2*, *Acot1*, *Acot2*, *Pdk4*, *Acaa1b*, *Cpt1*, *Fabp1*, and *Acadl*) were significantly up-regulated in the intestine of both A/J (obesity-resistant) and C57BL/6J (obesity-prone) mice fed with high fat diet. They observed that in the liver of A/J mice, high fat feeding significantly decreased the expression of *Mod1* and *Cyp4a10*. Also, the expression of *Mod1*, *Hmgcs2*, *Acot2* and *Pdk4* was not increased during high fat feeding in the muscle and white adipose tissue. We found that: (I) *Fabp1* was up-regulated and linearly responsive in all the three sections of the small intestine; (II) *Hmgcs2*, *Acot1*, *Acaa1b* and *Acadl* were linearly responsive and up-regulated in the proximal and middle sections; (III) in the proximal and middle sections, the expression of *Pdk4* was up-regulated and displayed a logarithmic response; (IV) *Acot2* was up-regulated and linearly responsive in the proximal section only. Al-Dwairi et al. (2014) reported that during diet-

induced obesity, ME1 over-expression in small intestine promoted the expression of hepatic genes associated with lipogenesis, cholesterol synthesis and cholesterol uptake. This suggests gatekeeper functionality of the small intestine, with changes in the expression of ME1 that influence metabolic processes in the liver. In our study, the gene *Me1* was up-regulated and logarithmically responsive in the proximal section, while the up-regulation was associated with linear response in the middle and distal sections. Moreover, we found that processes related to cholesterol homeostasis were particularly active in the middle small intestine and only highest fat intake impaired cholesterol transport and efflux (with a key role played by the down-regulation of ATP binding cassette transporters). In the small intestine, de Wit et al. reported that high fat modulates the expression of secreted proteins such as *Il18*, *Fgf15*, *Mif*, *Igfbp3* and *Angptl4* (de Wit et al. 2008). They suggested that these signaling molecules might have metabolic effects in liver, muscle and adipose tissue that underlie the development of the metabolic syndrome. In our study, the gene expression of *Il18* was down-regulated (changing from logarithmic response in the proximal to linear response in the middle section), while *Igfbp3* was linearly up-regulated in the middle section only.

3.5 Conclusions

Our approach to modelling nonlinear transcriptomic signatures in the small intestine revealed that - for a range of biological processes including cholesterol transport and homeostasis - there exists a 'tipping point' for fat intake, beyond which the relationship between fat intake and gene expression either weakens (logarithmic curves) or strengthens (exponential curves). For example, the shift from linear to exponential response observed for down-regulated ABC transporters (from the proximal to the middle section; see Table 3.11) is representative of the intense absorption capability of the middle section (i.e., only

extreme levels of fat intake down-regulate ABC transporters, while no effect is detected for fat intake ≤ 30E%).

Section	Response	Genes	GOBP
Proximal	Linear	*Cd36*	Cholesterol transport
		Abcg5; Abcg8	Cholesterol efflux Cholesterol homeostasis
Middle	Exponential	*Abcg1*	Cholesterol transport
Proximal - Middle	Linear - Logarithm	*Pla2g10*	Cholesterol homeostasis
Proximal - Middle	Linear - Exponential	*Abca1*	Cholesterol transport
Middle - Distal	Linear - Exponential	*Birc5*	Negative regulation of caspase activity
Proximal - Middle - Distal	Linear - Linear - Linear	*Serpina1b*	Acute-phase response

Table 3.11: Fat-responsive differentially expressed genes and the associated Gene Ontology Biological Process (GOBP) terms. Red colored genes: up-regulated; yellow colored genes: down-regulated. For each gene, site of expression (i.e., Section) and response type/trend (i.e., Response) are shown. These eight genes are the most representative for acute-phase response (and inflammatory processes, in all three intestine sections), cholesterol-related processes (i.e., cholesterol homeostasis, transport and efflux: in proximal and middle section) and negative regulation of caspase (in middle and distal section).

Our results highlight the relevance of nonlinear analysis for modelling more precisely the effects of diet on molecular activity in the small intestine. The observation of increasing relative importance of exponential responses and decreasing relative importance of logarithmic responses from the proximal to distal section is in agreement with the hypothesis that when the absorptive capacity of the intestinal epithelia is overloaded, the remaining fat will overflow to more distal sections (i.e., section-specific behavior). Although most of the genes related to inflammatory processes preserved their linear expression pattern along the whole small intestine (i.e., gene-specific behavior; e.g., see *Serpina1b* and *Cd36*), fat intake regulated in a section-specific fashion the transcriptomic response of *Birc5* (Table 3.11). The regionalized behavior of *Birc5* suggests that antiapoptotic mechanisms are particularly relevant in the middle section of small intestine and can represent an adaptation to counteract the lipotoxic effect of high fat diets on

intestinal cells (Desmarchelier et al. 2012; Gniuli et al. 2008). Future studies should assess whether the nonlinear patterns observed here are influenced by other factors (e.g., microbiome composition), and study the relevance of nonlinear responses in other tissues and clinically relevant markers.

Chapter 4

High fat perturbed pathways along the gut axis of C57BL/6J mice

4.1 Introduction

Dietary fat is an energy rich nutrient. Approximately 95% of dietary lipids in human diet are triacylglycerols that primarily contain long-chain fatty acids (with carbon number 16 or above), while the remaining are phospholipids (4.5%) and sterols (Niot et al. 2009). Fatty acids are essential components of the cell membrane, influence the transport activity across cell membranes (i.e., by influencing receptor and channel function), and modulate gene expression (Duplus et al. 2000). However, chronic fat overconsumption has been shown to increase the risk of a wide range of diseases including metabolic disorders. Even though the activity of the small intestine is crucial for fat absorption and digestion, its role is often underappreciated in the study of nutritional health and diseases (Niot et al. 2009).

The small intestine is a continuously proliferating epithelium with complex geographic differences in gene expression (Figure 1.1a). This is due to the fact that the structure of the intestinal epithelium differs along the cephalocaudal (duodenum to ileum) and crypt (undifferentiated) to villus (differentiated) axes (Gordon 1989). Peculiar differences in gene expression characterize each of the three small intestine axes (Figure 1.1b). Thus, the

number of expressed genes (or level of expression) is affected both by the epithelial cell lineage and by the location with respect to the three axes (Dusing et al. 2000; Gordon 1989). Murine small intestine shows adaptable cell proliferation and gene regulation in response to the dietary fat content (Niot et al. 2009). Such lipid-mediated intestinal adaptation is double edged with positive effects during food scarcity, whereas the risks and the progression of metabolic related diseases are maximized during food surplus (Figure 1.1d).

Considering the intricate architecture of the small intestine, we hypothesized three kinds of transcriptomic response to low- and high-fat diet intervention in mice. We characterized the differentially expressed genes intro linear and nonlinear (i.e., logarithmic and quadratic) response, as a function of the interaction between diet and 10 sections of the small intestine. Since dietary fat is both an important metabolic fuel and an integral part of the cell membrane, we examined diet-responsive metabolic and transport processes that were enriched from the proximal to the distal section of small intestine. After having identified the lists of significantly over-represented biological processes, we assessed the predominance of specific transcriptomic responses (i.e., either linear or nonlinear functions) with respect to either metabolic or transport processes. Since our main interest was related to the investigation of regional differences, we checked whether the 10 sections of small intestine were commonly affected by the fat content (i.e., showing gene-specific responses) or displayed peculiar transcriptomic patterns (i.e., section-specific responses).

4.2 Materials and methods

4.2.1 Microarray data

We analyzed microarray transcriptomic data from the intestinal mucosa of C57BL/6J male mice in 10 sections. The mice were fed with three kinds of diet for two weeks: high fat

(HF), low fat (LF) and chow (see Table 2.2 for dietary composition). The data contained 21,266 genes and 120 arrays (n = 4 per site, but 4 arrays were removed during the phases of data preprocessing and quality assessment, and therefore we analyzed 116 arrays). The Affymetrix mouse array used for this study is mogene11stmmentrezgprobe.

4.2.2 Data pre-processing

We used affyPLM (Bolstad 2004) library in R (Gentleman et al. 2004; R Core Team 2015) for data preprocessing and quality assessment. Background correction was performed using RMA (Robust Multichip Average; see Irizarry et al. 2003), whereas data normalization was achieved by using the quantile (Bolstad 2001) and UPC (Universal Probability of expression Codes) function of the SCAN.UPC library (Piccolo et al. 2013). Non-specific filtering of the genes whose intensity was above 0.50 (which corresponded to UPC output values ranging from 0 to 1) in at least four arrays were chosen for further analysis (i.e., differential gene expression analysis) with the genefilter library (Gentleman et al. 2012).

4.2.3 Differential gene expression analysis

To find differentially expressed genes (DEG) through the small intestine while using the resolution of 10 sections, we have modelled gene expression as a function of the interaction between diet and site (using the limma library; see Smyth 2005).The independent variable diet was considered as a categorical parameter (i.e., HF, LF, and Chow). For modelling the geographical location along the longitudinal axis of small intestine (from the proximal to the distal section) we assigned continuous numerical values from 1 to 10 to the independent variable called site. We fitted linear and nonlinear (i.e., logarithmic and quadratic) responses to model gene expression as a function of: (1) the level of fat intake; and (2) the small intestine section.

Figure 4.1 illustrates three theoretical examples (i.e., for the linear, logarithmic and quadratic response types) where the expression of an up-regulated gene is described either with linear or nonlinear response, as a function of the intestinal sections. We used Benjamini and Hochberg's false discovery rate (1995) for multiple test correction and the significance threshold was 0.1(adjusted p-value = 0.1). After having found which genes responded in a significant way to linear and nonlinear response types (i.e., linear, logarithmic and quadratic), we obtained diet responsive DEG by applying contrast tests. Contrast tests have been carried out between: (1) HF vs. LF; (2) LF vs. Chow. When a significant DEG was associated with more than one response, we selected the response with the lowest adjusted p-value.

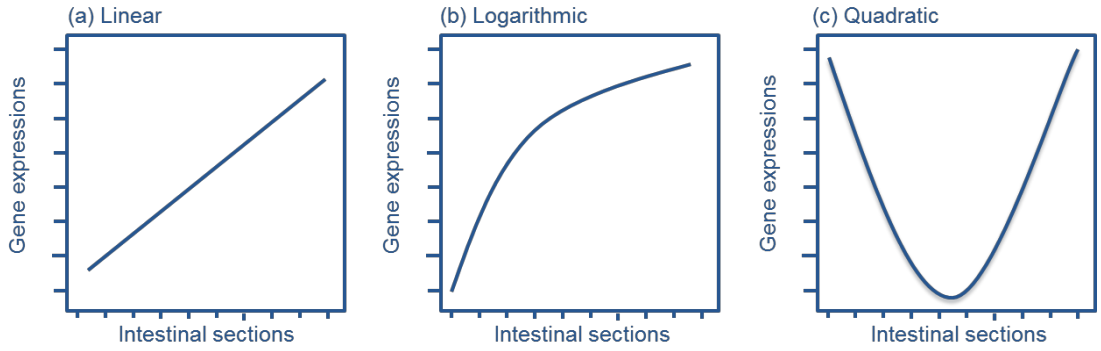

Figure 4.1: Theoretical response types. (a) The linear response portrays gradual increase in expression from the proximal to the distal section. (b) The logarithmic response represents a saturation behavior that is particularly pronounced along the most downstream sections. (c) The quadratic response describes major contribution of the proximal and distal sections in terms of gene expression. The response types are illustrated for three hypothetical, up-regulated genes.

4.2.4 Weighted co-expression network analysis

We performed hierarchical clustering by using the mean expression value of the 1,650 DEG in the 30 samples (i.e., 10 samples per dietary interventions). This served to check associations (and similarities) in gene expression across samples collected from the ten intestinal sections. Using the WGCNA (Weighted correlation network analysis; see Langfelder & Horvath 2008) library we defined clusters (i.e., modules) of highly correlated DEG and investigated their relation with the 10 intestinal sections (i.e., to see whether these modules responded in a coherent way along the longitudinal axis of small intestine). Then, gene set enrichment test was carried out to detect over-represented Gene Ontology, Biological Processes in each module separately.

4.2.5 Gene set enrichment test

We performed Gene Ontology Biological Process (GOBP) enrichment test to functionally annotate each of the modules that were significantly correlated with the small intestine longitudinal axis (we selected the modules that displayed a correlation ≥ 0.85 with the 10

sections; p-value threshold = 0.1); gene set enrichment was tested using the gage library (Generally Applicable Gene-set Enrichment; see Luo et al. 2009). Three kinds of analysis were performed per module by: (1) ignoring the diet effect (abbreviated as ND that stands for non-diet); (2) focusing on HF responsive processes (HF vs. LF); (3) studying LF responsive processes (LF vs. Chow). For each analysis, three lists of over-represented GOBP terms were obtained, according to the direction of regulation of DEG: (1) processes composed of up-regulated genes only; (2) processes including down-regulated genes only; (3) processes that were enriched in genes regulated in both directions (i.e., both up- and down-regulation). Within each module, GOBP that displayed two regulatory directions were classified by giving preference to the lowest q-value. Then unique GOBP terms described by specific regulatory directions were obtained. To reduce the complexity, each GOBP output was clustered based on the pairwise GO-term similarity using the GOSim library (Fröhlich et al. 2007). Since we aimed at finding dietary-fat perturbed biological functions along the gut axis, we focused our subsequent analysis on clusters that were relevant to metabolism and transport processes only.

4.2.6 KEGG disease pathway enrichment test

Using gage library (Luo et al. 2009), KEGG (Kyoto Encyclopedia of Genes and Genomes) disease pathway enrichment tests were performed for the modules that presented significant correlation with the ten intestinal sections. Three kinds of analysis were performed: (1) ND; (2) HF vs. LF; and (3) LF vs. Chow. For each analysis, outputs were classified in three groups of KEGG disease pathways that differ in regulatory direction of the DEG: (1) pathways with up-regulated genes only; (2) pathways with down-regulated genes only; and, (3) pathways consisting of both up- and down-regulated genes.

4.2.7 Dominant response-type and intestinal section

For the significant GOBP terms found in each cluster (i.e., each cluster was characterized separately for transport and metabolic processes), we considered the response with the highest value of gene count as the dominant. The section (i.e., within the first to the tenth section) that has the number of differentially expressed genes (i.e., associated to a specific GOBP term) either highest (termed as Max Site) or lowest (termed as Min Site) was considered as being actively or passively responsive to diet, respectively.

4.3 Results

4.3.1 Diet responsive DEG

After data pre-processing and quality assessment we obtained a set of 9,432 transcriptionally active genes. We found 1,650 diet responsive DEG. The final count of total DEG (as well as the count of DEG for each response type) in the HF vs. LF contrast was considerably higher than in the LF vs. Chow contrast. The majority of the HF responsive DEG belonged to the logarithmic response (Figure 4.2).

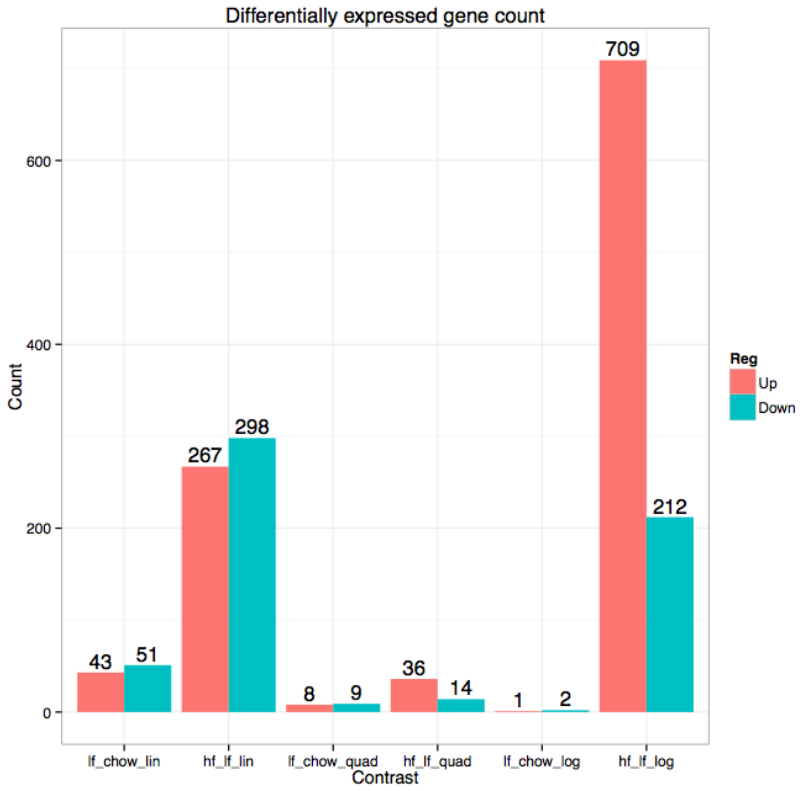

Figure 4.2: Up- and down-regulated DEG. We obtained diet responsive DEG and response types using the limma library (adjusted p-value = 0.1). Red and green bars represent up- and down-regulated counts of DEG, respectively. For each pair of bars, the x-axis label is written with the following format: experimental diet, control diet and response type. Diets: lf = LF; hf = HF; response types: lin = linear; log = logarithmic; quad = quadratic.

4.3.2 Significant modules

The hierarchical clustering of the 30 samples was defined by relying on the mean expression value of 1,650 DEG in 10 samples per dietary intervention. Such clustering suggested high similarities in the site 1 of the small intestine in case of Chow and LF diet while a different behavior was found in presence of HF diet. However, site 1 showed higher similarities among each of the three dietary interventions than what could be observed with the more downstream parts (Figure 4.3). Section 2 through 6 in HF was more similar to section 2 to 4 in LF and Chow. Section 7 through 9 in HF and LF differed slightly from Chow.

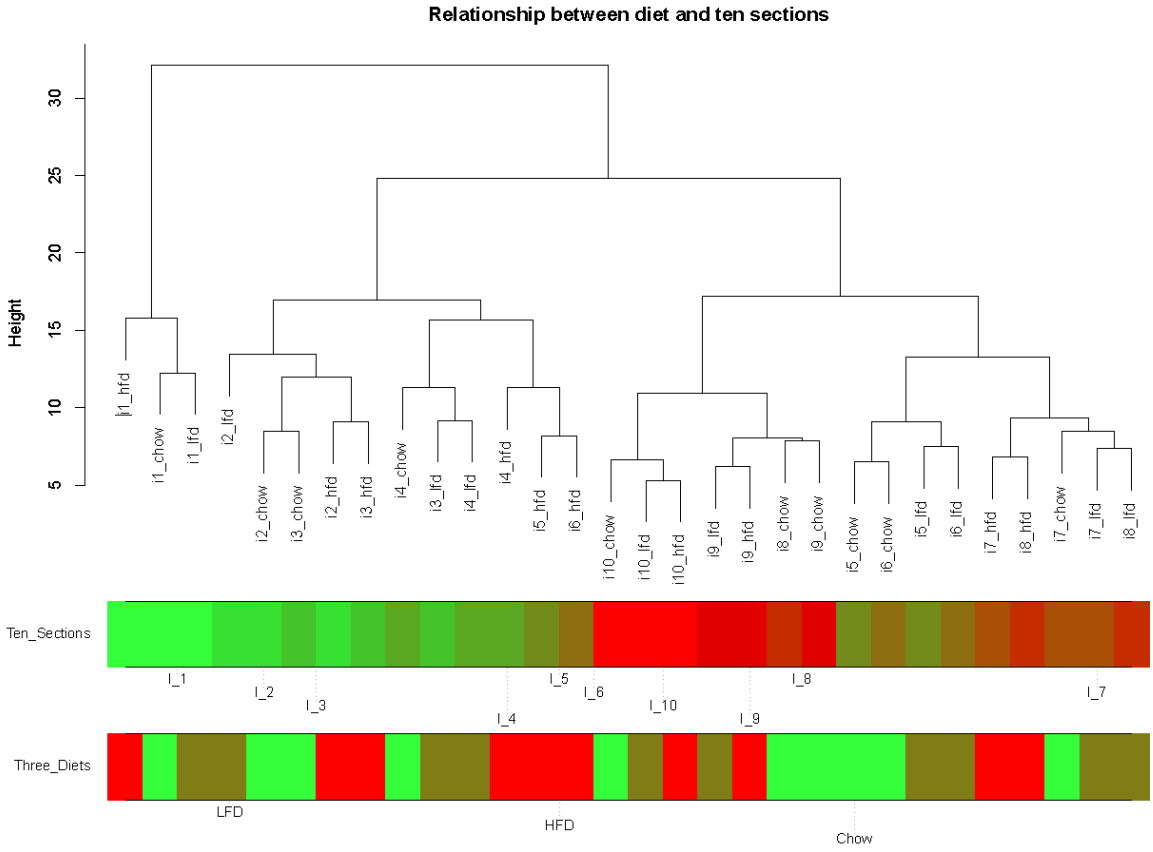

Figure 4.3: Hierarchical clustering of transcriptomic data of ten samples per dietary intervention. Ten samples are considered for each diet: high-fat – HFD; low-fat – LFD; and chow. Each sample is labelled as intestinal section number followed by diet type. For example, sample collected from first section of the intestine with the intervention of HFD was labelled as, "i1_hfd". The relationships between diet and ten samples were further illustrated with heat maps (i.e., for ten sections – Ten_Sections; and for dietary interventions – Three_Diets) below the dendrogram.

As implemented in the WGCNA library (Langfelder & Horvath 2008), the Pearson correlation between eigengene values of each module and the ten intestinal sections (i.e., continuous numerical value ranging from 1 through 10) was calculated, including the corresponding p-values. The threshold for correlation is > 0.7. Since the correlation values of modules blue, brown and green were larger than 0.8 (and associated p-values significant), they were considered as highly correlated (Figure 4.4). The three significant modules were: (1) blue, which was positively correlated (i.e., 0.89) with the longitudinal axis of small intestine; (2) brown, which was negatively correlated (i.e., -0.91) with the 10

intestinal sections; (3) green, which showed a negative correlation (i.e., -0.92) along the small intestine. The majority of the genes were HF-responsive and HF vs. LF contrast had the highest count of DEG in comparison with LF vs. Chow (Table 4.1). The highest number of DEG in both the blue and the green module showed logarithmic response as a function of the small intestine site whereas the brown module had the highest number of genes responding in a linear way (Figure 4.5).

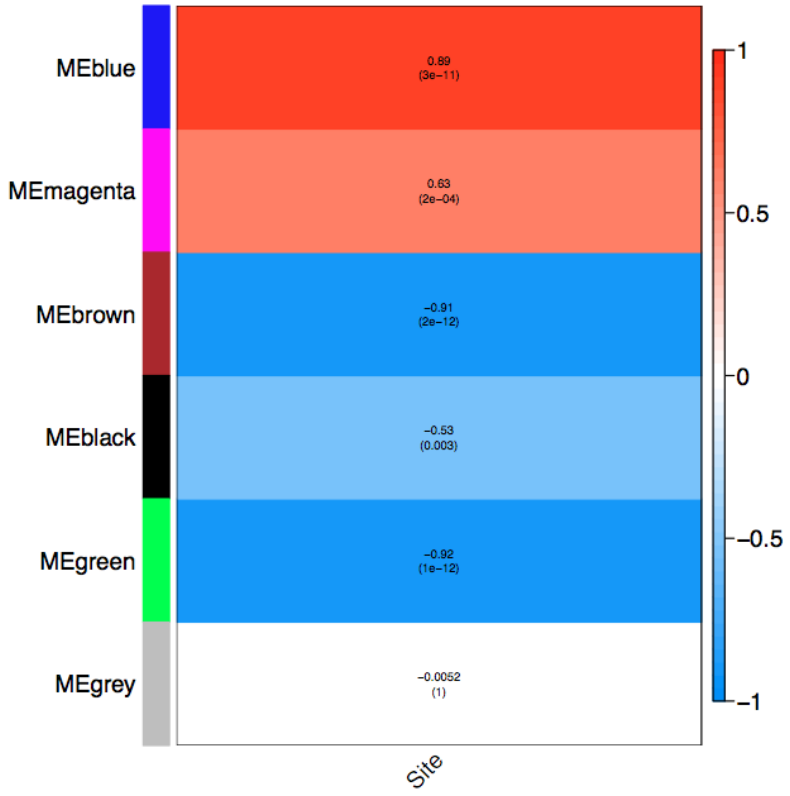

Figure 4.4: Associations between module eigengene and ten intestinal sections. The vertical axis represents eigengene profile of each module (i.e., ME for module eigengene followed by the color code of each module; e.g., MEblue – color cording is used to name the modules but has no biological meaning). The color in each cell corresponds to the correlation (i.e., between module eigengene and the ten intestinal sections); color legend is on the right side. The correlation (upper value) and p-value (bottom value within parenthesis) between module eigengene and the ten intestinal sections are displayed within each cell.

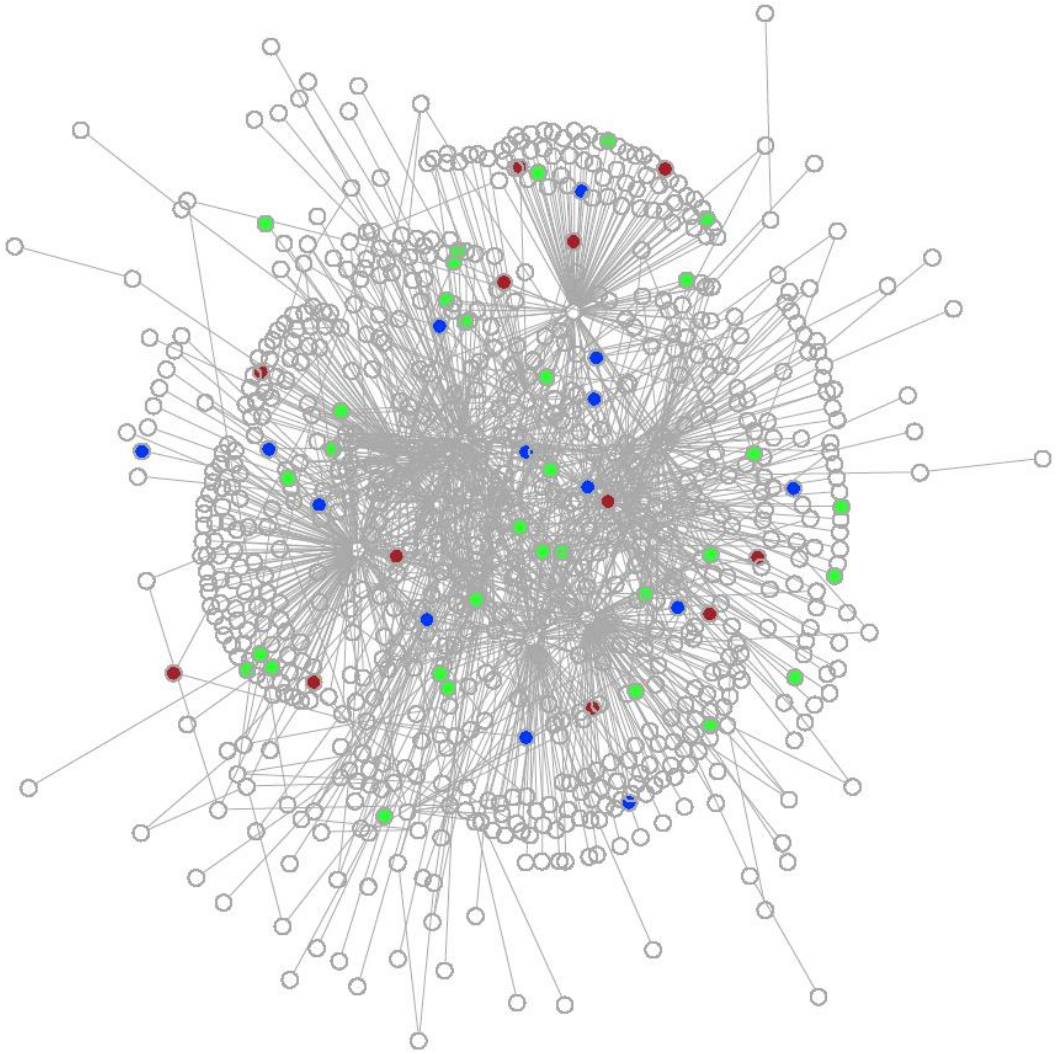

Figure 4.5 Genes that respond significantly to diet and ten intestinal sections. Using weighted co-expression network analysis, three main modules were identified: (1) the green module contains genes involved in metabolic and transport processes that are modulated by HF, with a prevalence of logarithmic responses; (2) the brown module includes a predominant group of DEGs that respond in a linear way to HF; (3) logarithmic responsive DEGs not enriched with any biological process related to dietary fat are in the blue module. For this illustration, the network was constructed using the protein interactions from Biogrid (Stark et al. 2011). First, the giant component (i.e., the connected sub-network that includes the largest number of proteins) was extracted and used as a scaffold to show the genes identified with weighted co-expression network analysis. Second, topological communities were found using the "spinglass.community" function from the R package igraph (Csardi & Nepusz 2006). Third, the topological community most significantly enriched with genes from the green, brown and blue modules (i.e., modules that were previously characterized with weighted co-expression network analysis) was visualized and nodes color coded according to the color of the module (i.e., blue, brown and green).

Contrasts	Response	Green	Brown	Blue
HF vs. LF	Linear	250	85	33
	Quadratic	29	5	0
	Logarithmic	262	66	147
LF vs. Chow	Linear	31	27	3
	Quadratic	7	1	0
	Logarithmic	0	2	1
Total		579	186	184

Table 4.1: DEG counts per module. We reported the counts of DEG concerning three modules (Green, Brown, and Blue); they refer to genes responsive to HF (HF vs. LF) and LF (LF vs. Chow) diets. Modules were extracted using the WGCNA library. Red colored numbers represent the highest counts in each column.

4.3.3 Enriched GOBP

The number of enriched GOBP terms differed considerably among the three modules (Table 4.2). The green module had the most diet-responsive GOBP terms and was followed by the brown, whereas the majority of GOBP terms in the blue module belonged to ND (i.e., non-diet) processes.

Analysis	Regulation	Green	Brown	Blue
ND	Up	-	-	-
	Down	142	27	11
	Both	160	28	24
HF vs. LF	Up	11	5	-
	Down	187	8	-
	Both	58	11	-
LF vs. Chow	Up	120	3	-
	Down	8	-	10
	Both	35	14	-

Table 4.2: Count of significant GOBP terms per module. The enrichment test was performed using the gage library and refers to three modules (Green, Brown, and Blue). Three analyses (ND - non-diet, HF responsive, and LF responsive) were conducted to find diet-responsive and diet-independent GOBP terms. Three outputs were obtained from each analysis: Up – up-regulated processes; Down – down-regulated processes; and Both – processes that show regulatory patterns involving both up- and down-regulated genes.

We analyzed GOBP terms in clusters based on their pairwise GO-term similarity (i.e., the enriched GOBP terms were categorized into clusters). One of our goals was to find dietary-fat perturbed biological processes along the gut axis. Therefore, we considered only GOBP clusters that consisted of metabolic and transport processes (separately for the green the brown modules; see Table 4.3). Our main focus was on the green module as it contained the highest number of diet-responsive GOBP terms and DEG (Table 4.3). Moreover, the green module was the one with the most significant and strongest correlation with the ten sections of the small intestine (Figure 4.4).

Modules	Analysis	Regulation	Count					
			Clusters	Transport processes	DEG	Clusters	Metabolic Processes	DEG
Green	HF vs. LF	Up	1	2	35	1	8	63
		Down	3	6	40	12	46	165
		Both	2	3	69	7	13	82
	LF vs. Chow	Up	2	4	40	8	26	226
		Down	1	1	35	1	4	26
		Both	1	2	70	3	9	65
Brown	HF vs. LF	Up	-	-	-	3	5	45
		Down	-	-	-	2	6	78
		Both	1	5	25	1	3	17
	LF vs. Chow	Up	1	3	11	-	-	-
		Down	-	-	-	-	-	-
		Both	1	3	19	2	6	49

Table 4.3 Count of significant GOBP terms and the corresponding DEG per module. Within each module (i.e., Green and Brown), per analysis group (for example, HF vs. LF), clusters of GOBP terms representing transport and metabolic processes were considered. The count of GOBP terms were reported separately for up- and down-regulated processes.

In case of HF diet, the green module showed over-representation of processes related to lipid metabolism and transport (Table 4.4). Other processes that are not related to lipid metabolism were down-regulated (Table 4.4). The response of the green module to LF was the opposite if compared to the response in presence of HF diet: we observed up-regulation of various processes related to carbohydrate metabolism (Table 4.4) and under-representation of lipid metabolism and transport processes (Table 4.4).

Diet	Clus ID	Type	Regulation	GOBP	Qval	ICval	count	Response:count
HF	1	TP	Up	GO:0016192 vesicle-mediated transport	<0.01	4.51	35	Hf-lin:13; Hf-log:15; Lf-lin:6; Lf-quad:1
				GO:0006897 endocytosis	0.06	5.38	14	Hf-lin:7; Hf-log:6; Lf-quad:1
HF	2	MP	Up	GO:0044242 cellular lipid catabolic process	0.01	6.62	12	Hf-lin:6; Hf-log:5; Hf-quad:1
				GO:0006631 fatty acid metabolic process	0.01	5.41	27	Hf-lin:11; Hf-log:12; Hf-quad:2; Lf-lin:2
				GO:0006644 phospholipid metabolic process	0.01	5.86	14	Hf-lin:5; Hf-log:9
				GO:0016042 lipid catabolic process	0.01	5.97	14	Hf-lin:6; Hf-log:7; Hf-quad:1
				GO:0045834 positive regulation of lipid metabolic process	0.03	6.91	10	Hf-lin:7; Hf-log:1; Hf-quad:1; Lf-lin:1
				GO:0046486 glycerolipid metabolic process	0.03	5.89	15	Hf-lin:8; Hf-log:7
				GO:0032787 monocarboxylic acid metabolic process	0.06	5.17	32	Hf-lin:12; Hf-log:15; Hf-quad:2; Lf-lin:3
				GO:0046434 organophosphate catabolic process	0.07	8.99	24	Hf-lin:10; Hf-log:11; Hf-quad:2; Lf-lin:1
HF	48	TP	Down	GO:0055085 transmembrane transport	0.02	3.98	31	Hf-lin:13; Hf-log:15; Hf-quad:3
				GO:0034220 ion transmembrane transport	0.03	4.50	15	Hf-lin:3, Hf-log:9; Hf-quad:3
	49			GO:0006812 cation transport	0.03	4.24	17	Hf-lin:3; Hf-log:10; Hf-quad:3; Lf-lin:1
				GO:0006811 ion transport	0.03	3.56	30	Hf-lin9; Hf-log16; Hf-quad:4; Lf-lin:1
				GO:0030001 metal ion transport	0.03	4.46	11	Hf-lin:3; Hf-log:5; Hf-quad:3
	61			GO:0043269 regulation of ion transport	0.03	5.41	10	Hf-lin2; Hf-log:5; Hf-quad:2; Lf-lin:1
HF	11	MP	Down	GO:0010557 positive regulation of macromolecule biosynthetic process	<0.01	4.04	32	Hf-lin20; Hf-log:8; Hf-quad:3; Lf-lin:1
				GO:0010628 positive regulation of gene expression	0.01	4.11	27	Hf-lin:15; Hf-log:8; Hf-quad:3; Lf-lin:1

				GO:0010604 positive regulation of macromolecule metabolic process	0.01	3.68	54	Hf-lin:30; Hf-log:18; Hf-quad:3; Lf-lin:3
				GO:0045944 positive regulation of transcription from RNA polymerase II promoter	0.01	4.71	22	Hf-lin:12; Hf-log:6; Hf-quad:3; Lf-lin:1
	20			GO:0045935 positive regulation of nucleobase-containing compound metabolic process	<0.01	4.08	32	Hf-lin:18; Hf-log:10; Hf-quad3; Lf-lin:1
				GO:0045893 positive regulation of transcription, DNA-dependent	0.01	4.19	27	Hf-lin:15; Hf-log:8; Hf-quad:3; Lf-lin:1
				GO:0051254 positive regulation of RNA metabolic process	0.01	4.17	28	Hf-lin16; Hf-log:8; Hf-quad:3; Lf-lin:1
				GO:0051173 positive regulation of nitrogen compound metabolic process	0.01	4.06	34	Hf-lin:18; Hf-log:11; Hf-quad:3; Lf-lin:2
	25			GO:0005996 monosaccharide metabolic process	<0.01	5.64	23	Hf-lin:8; Hf-log:11; Hf-quad:2; Lf-lin:2
				GO:0019318 hexose metabolic process	0.01	5.85	22	Hf-lin8; Hf-log:11; Hf-quad:2; Lf-lin:1
				GO:0006006 glucose metabolic process	0.01	6.08	18	Hf-lin:6; Hf-log:10; Hf-quad:1; Lf-lin:1
				GO:0044262 cellular carbohydrate metabolic process	0.07	5.05	16	Hf-lin:7; Hf-log:6; Hf-quad:1; Lf-lin:2
	29			GO:0010605 negative regulation of macromolecule metabolic process	0.01	4.12	36	Hf-lin:21; Hf-log:12; Hf-quad:2; Lf-lin:1
				GO:0031324 negative regulation of cellular metabolic process	0.01	4.16	39	Hf-lin:22; Hf-log:13; Hf-quad:3; Lf-lin:1
				GO:0010629 negative regulation of gene expression	0.05	4.39	23	Hf-lin:14; Hf-log:7; Hf-quad:2
				GO:2000113 negative regulation of cellular macromolecule biosynthetic process	0.05	4.45	23	Hf-lin:13; Hf-log:8; Hf-quad:2
				GO:0010558 negative regulation of macromolecule biosynthetic process	0.06	4.41	26	Hf-lin:15; Hf-log:9; Hf-quad:2
				GO:0031327 negative regulation of cellular biosynthetic process	0.1	4.37	27	Hf-lin15; Hf-log:9; Hf-quad:3
	30			GO:0031399 regulation of protein modification process	0.01	4.42	40	Hf-lin:21; Hf-log:15; Hf-quad:1; Lf-lin:3

				GO:0051246 regulation of protein metabolic process	0.04	3.99	48	Hf-lin:26; Hf-log:18; Hf-quad:1; Lf-lin:3
				GO:0051247 positive regulation of protein metabolic process	0.05	5.02	30	Hf-lin:17; Hf-log:10; lf_chow_lin3
				GO:0031401 positive regulation of protein modification process	0.06	5.48	28	Hf-lin:15; Hf-log:10; Lf-lin:3
				GO:0032268 regulation of cellular protein metabolic process	0.06	4.17	46	Hf-lin:25; Hf-log:17; Hf-quad:1; Lf-lin3
				GO:0032270 positive regulation of cellular protein metabolic process	0.08	5.25	29	Hf-lin:16; Hf-log:10; Lf-lin:3
	33			GO:0005975 carbohydrate metabolic process	0.01	4.69	39	Hf-lin:12; Hf-log:22; Hf-quad:2; Lf-lin:3
				GO:0016051 carbohydrate biosynthetic process	0.05	6.38	10	Hf-lin:4; Hf-log:5; Lf-lin:1
	36			GO:0060255 regulation of macromolecule metabolic process	0.01	2.56	101	Hf-lin:57; Hf-log:37; Hf-quad:3; Lf-lin:4
				GO:0006355 regulation of transcription. DNA-dependent	0.02	3.04	61	Hf-lin:36; Hf-log:21; Hf-quad:3; Lf-lin:1
				GO:0010556 regulation of macromolecule biosynthetic process	0.03	2.93	73	Hf-lin:43; Hf-log:26; Hf-quad:3; Lf-lin:1
				GO:2000112 regulation of cellular macromolecule biosynthetic process	0.04	2.97	69	Hf-lin:41; Hf-log:24; Hf-quad:3; Lf-lin:1
				GO:0010468 regulation of gene expression	0.05	2.86	71	Hf-lin:42; Hf-log:25; Hf-quad:3; Lf-lin:1
	45			GO:0051252 regulation of RNA metabolic process	0.02	3.01	61	Hf-lin:36; Hf-log:21; Hf-quad:3; Lf-lin:1
				GO:0006351 transcription, DNA-dependent	0.02	2.72	62	Hf-lin:37; Hf-log:21; Hf-quad:3; Lf-lin:1
				GO:0032774 RNA biosynthetic process	0.02	2.72	62	Hf-lin:37; Hf-log:21; Hf-quad:3; Lf-lin:1
				GO:0019219 regulation of nucleobase-containing compound metabolic process	0.07	2.86	71	Hf-lin:42; Hf-log:24; Hf-quad:3; Lf-lin:2
	70			GO:0045934 negative regulation of nucleobase-containing compound metabolic process	0.05	4.45	23	Hf-lin:14; Hf-log:7; Hf-quad:2
				GO:0051172 negative regulation of nitrogen compound metabolic	0.05	4.44	23	Hf-lin:14; Hf-log:7; Hf-quad:2

				process				
				GO:0051253 negative regulation of RNA metabolic process	0.07	4.51	22	Hf-lin:13; Hf-log:7; Hf-quad:2
				GO:0045892 negative regulation of transcription, DNA-dependent	0.07	4.56	21	Hf-lin:12; Hf-log:7; Hf-quad:2
				GO:0000122 negative regulation of transcription from RNA polymerase II promoter	0.09	5.17	12	Hf-lin:6; Hf-log:4; Hf-quad:2
	71			GO:0001933 negative regulation of protein phosphorylation	0.05	7.24	10	Hf-lin:5; Hf-log:4; Hf-quad:1
				GO:0010563 negative regulation of phosphorus metabolic process	0.05	6.79	10	Hf-lin:5; Hf-log:4; Hf-quad:1
				GO:0042326 negative regulation of phosphorylation	0.05	6.83	10	Hf-lin:5; Hf-log:4; Hf-quad:1
				GO:0045936 negative regulation of phosphate metabolic process	0.05	6.79	10	Hf-lin:5; Hf-log:4; Hf-quad:1
	38			GO:0044723 single-organism carbohydrate metabolic process	0.01	NA	30	Hf-lin:10; Hf-log:16; Hf-quad:2; Lf-lin:2
	73			GO:1901135 carbohydrate derivative metabolic process	0.07	NA	45	Hf-lin:15; Hf-log:24; Hf-quad:3; Lf-lin:2; Lf-quad:1
HF	10	TP	Both	GO:0006869 lipid transport	<0.01	5.87	10	Hf-lin:7; Hf-log:2; Hf-quad:1
				GO:0010876 lipid localization	0.02	5.70	12	Hf-lin:8; Hf-log:3; Hf-quad:1
	14			GO:0071702 organic substance transport	0.03	4.70	68	Hf-lin:33; Hf-log:23; Hf-quad:3; Lf-lin:7; Lf-quad:2
HF	1	MP	Both	GO:0008202 steroid metabolic process	<0.01	5.46	10	Hf-lin:4; Hf-log:4; Hf-quad:1; Lf-lin:1
				GO:0044255 cellular lipid metabolic process	<0.01	4.50	45	Hf-lin:17; Hf-log:22; Hf-quad:3; Lf-lin:3
				GO:0006629 lipid metabolic process	<0.01	3.92	54	Hf-lin:19; Hf-log:27; Hf-quad:4; Lf-lin:4
				GO:0008610 lipid biosynthetic process	0.03	4.89	33	Hf-lin:15; Hf-log:14; Hf-quad:2; Lf-lin:2
	4			GO:0044283 small molecule biosynthetic process	<0.01	4.65	30	Hf-lin:13; Hf-log:12; Hf-quad:2; Lf-lin:3

				GO:0006633 fatty acid biosynthetic process	0.08	6.32	12	Hf-lin:5; Hf-log:5; Hf-quad:1; lf-lin:1
	6			GO:0006066 alcohol metabolic process	<0.01	4.91	19	Hf-lin:9; Hf-log:7; Hf-quad:1; lf-lin:2
	8			GO:0010565 regulation of cellular ketone metabolic process	<0.01	7.14	11	Hf-lin:7; Hf-log:2; Hf-quad:1; lf-lin:1
				GO:0042180 cellular ketone metabolic process	<0.01	4.39	11	Hf-lin:7; Hf-log:2; Hf-quad:1; lf-lin:1
	9			GO:0019216 regulation of lipid metabolic process	<0.01	6.11	15	Hf-lin:8; Hf-log:4; Hf-quad:2; lf-lin:1
				GO:0046890 regulation of lipid biosynthetic process	<0.01	6.94	12	Hf-lin:7; Hf-log:2; Hf-quad:2; lf-lin:1
	29			GO:0033993 response to lipid	0.1	8.22	18	Hf-lin:11; Hf-log:5; Hf-quad:1; lf-quad:1
	2			GO:1901615 organic hydroxy compound metabolic process	<0.01	NA	23	Hf-lin:11; Hf-log:7; Hf-quad:1; lf-lin:4
LF	29	TP	UP	GO:0055085 transmembrane transport	0.01	3.98	31	Hf-lin:13; Hf-log:15; Hf-quad:3
	35			GO:0006812 cation transport	0.04	4.24	17	Hf-lin:3; Hf-log:10; Hf-quad:3; lf-lin:1
				GO:0006811 ion transport	0.06	3.56	30	Hf-lin:9; Hf-log:16; Hf-quad:4; lf-lin:1
				GO:0043269 regulation of ion transport	0.09	5.41	10	Hf-lin:2; Hf-log:5; Hf-quad:2; lf-lin:1
LF	1	MP	UP	GO:0005975 carbohydrate metabolic process	<0.01	4.69	39	Hf-lin:12; Hf-log:22; Hf-quad:2; lf-lin:3
				GO:0016051 carbohydrate biosynthetic process	<0.01	6.38	10	Hf-lin:4; Hf-log:5; lf-lin:1
				GO:0046390 ribose phosphate biosynthetic process	0.07	9.95	12	Hf-lin:2; Hf-log:9; Hf-quad:1
	2			GO:0005996 monosaccharide metabolic process	<0.01	5.65	23	Hf-lin:8; Hf-log:11; Hf-quad:2; lf-lin:2
				GO:0019318 hexose metabolic process	<0.01	5.85	22	Hf-lin:8; Hf-log::11; Hf-quad:2; Lf-lin:1
				GO:0006006 glucose metabolic process	<0.01	6.08	18	Hf-lin:6; Hf-log:10; Hf-quad:1; Lf-lin:1
				GO:0044262 cellular carbohydrate metabolic process	<0.01	5.05	16	Hf-lin:7; Hf-log:6; Hf-quad:1; Lf-lin:2

				GO:0016052 carbohydrate catabolic process	0.01	6.59	12	Hf-lin:3; Hf-log:7; Hf-quad:1; Lf-lin:1
				GO:0019320 hexose catabolic process	0.01	7.10	12	Hf-lin:3; Hf-log:7; Hf-quad:1; Lf-lin:1
				GO:0046365 monosaccharide catabolic process	0.01	7.08	12	Hf-lin:3; Hf-log:7; Hf-quad:1; Lf-lin:1
	21			GO:0016311 dephosphorylation	<0.01	5.31	20	Hf-lin:9; Hf-log:8; Hf-quad:1; Lf-lin:1; Lf-quad:1
				GO:0006468 protein phosphorylation	0.02	3.95	46	Hf-lin:23; Hf-log:17; Hf-quad:2; Lf-lin:4
				GO:0016310 phosphorylation	0.02	3.56	70	Hf-lin:31; Hf-log:31; Hf-quad:2; Lf-lin:6
				GO:0006793 phosphorus metabolic process	0.09	3.40	125	Hf-lin:51; Hf-log:59; Hf-quad:6; Lf-lin:8; Lf-quad:1
	32			GO:0031324 negative regulation of cellular metabolic process	0.03	4.16	39	Hf-lin:22; Hf-log:13; Hf-quad:3; Lf-lin:1
				GO:0010605 negative regulation of macromolecule metabolic process	0.04	4.11	36	Hf-lin:21; Hf-log:12; Hf-quad:2; Lf-lin:1
				GO:0009892 negative regulation of metabolic process	0.06	4.02	43	Hf-lin:24; Hf-log:15; Hf-quad:3; Lf-lin:1
	34			GO:0019222 regulation of metabolic process	0.04	2.32	127	Hf-lin:65; Hf-log:47; Hf-quad:6; Lf-lin:7; Lf-quad:2
				GO:0006357 regulation of transcription from RNA polymerase II promoter	0.1	4.06	30	Hf-lin:17; Hf-log:9; Hf-quad:3; Lf-lin:1
	40			GO:0045944 positive regulation of transcription from RNA polymerase II promoter	0.07	4.71	22	Hf-lin:12; Hf-log:6; Hf-quad:3; Lf-lin:1
				GO:0051173 positive regulation of nitrogen compound metabolic process	0.09	4.06	34	Hf-lin:18; Hf-log:11; Hf-quad:3; Lf-lin:2
				GO:0051254 positive regulation of RNA metabolic process	0.09	4.17	28	Hf-lin:16; Hf-log:8; Hf-quad:3; Lf-lin:1
				GO:0010628 positive regulation of gene expression	0.09	4.11	27	Hf-lin:15; Hf-log:8; Hf-quad:3; Lf-lin:1

				GO:0045893 positive regulation of transcription. DNA-dependent	0.09	4.20	27	Hf-lin:15; Hf-log:8; Hf-quad:3; Lf-lin:1
	43			GO:0006520 cellular amino acid metabolic process	0.09	5.21	20	Hf-lin:10; Hf-log:6; Hf-quad:2; Lf-lin:2
	4			GO:0044723 single-organism carbohydrate metabolic process	<0.01	NA	30	Hf-lin:10; Hf-log:16; Hf-quad:2; Lf-lin:2
LF	3	TP	Down	GO:0016192 vesicle-mediated transport	0.01	4.51	35	Hf-lin:13; Hf-log:15; Lf-lin:6; Lf-quad:1
LF	2	MP	Down	GO:0044242 cellular lipid catabolic process	0.01	6.62	12	Hf-lin:6; Hf-log:5; Hf-quad:1
				GO:0006644 phospholipid metabolic process	0.01	5.86	14	Hf-lin:5; Hf-log:9
				GO:0016042 lipid catabolic process	0.01	5.97	14	Hf-lin:6; Hf-log:7; Hf-quad:1
				GO:0030258 lipid modification	0.06	6.62	10	Hf-lin:6; Hf-log:3; Hf-quad:1
LF	6	TP	Both	GO:0071702 organic substance transport	0.02	4.70	68	Hf-lin:33; Hf-log:23; Hf-quad:3; Lf-lin:7; 2Lf-quad:
				GO:0032940 secretion by cell	0.08	4.83	18	Hf-lin:8; Hf-log:8; Lf-lin:2
LF	2	MP	Both	GO:0008202 steroid metabolic process	0.01	5.46	10	Hf-lin:4; Hf-log:4; Hf-quad:1; Lf-lin:1
				GO:0044255 cellular lipid metabolic process	0.09	4.50	45	Hf-lin:17; Hf-log:22; Hf-quad:3; Lf-lin:3
	3			GO:0010565 regulation of cellular ketone metabolic process	0.01	7.14	11	Hf-lin:7; Hf-log:2; Hf-quad:1; Lf-lin:1
				GO:0042180 cellular ketone metabolic process	0.01	4.39	11	Hf-lin:7; Hf-log:2; Hf-quad:1; Lf-lin:1
				GO:0006066 alcohol metabolic process	0.03	4.91	19	Hf-lin:9; Hf-log:7; Hf-quad:1; Lf-lin:2
				GO:0044283 small molecule biosynthetic process	0.04	4.65	30	Hf-lin:13; Hf-log:12; Hf-quad:2; Lf-lin:3
	1			GO:1901615 organic hydroxy compound metabolic process	<0.01	NA	23	Hf-lin:11; Hf-log:7; Hf-quad:1; Lf-lin:4
				GO:1901617 organic hydroxy compound biosynthetic process	<0.01	NA	11	Hf-lin:5; Hf-log:3; Hf-quad:1; Lf-lin:2
				GO:0044711 single-organism	0,04	NA	30	Hf-lin:13; Hf-log:12;

				biosynthetic process				Hf-quad:2; Lf-lin:3

Table 4.4: Diet responsive GOBP terms. Two types of diet are considered: HF and LF. GOBP terms can be either up- or down-regulated (Regulation), and be associated to linear, logarithmic or quadratic response types. GOBP terms refer to either metabolic (MP) or transport processes (TP). Results were obtained using the gage library. GOBP terms with the same cluster ID (ClusID) belong to the same cluster (this was based on GO term similarity, measured using the GOSim library). For each GOBP term, the results include: the multiple test correction computed using the Benjamini and Hochberg false discovery rate (Qval); the information content value (ICval) that was calculated for each GO term using GOSim; the number of DEGs (count); the response specific count of the DEGs and the corresponding response name for each count of responses (Response:count). For example (in the first row), GO:0016192 vesicle-mediated transport contained 35 DEG (count), out of which 13 (Response:count) were HF and linear responsive (Hf- lin); the dominant response (in red color) represents the response with the highest gene count; e.g., in the first column, 15 (Response:count) out of 35 (count) genes were HF and logarithmic responsive (Hf-log).

4.3.4 KEGG disease pathways

Among the three modules, only one HF (i.e., HF vs. LF) responsive disease pathway was enriched. The over-represented term was pathways in cancer (mmu05200); it consists of genes that are both up- and down-regulated and was found in the green module (Table 4.5).

Disease KEGG pathway	p.geomean	stat.mean	p.val	q.val	set.size	genes
mmu05200: Pathways in cancer	0.16	1.04	<0.01	<0.01	12	*Adcy6*, *Bcl2*, *Bcl2l1*, *Bmp2*, *Csf2ra*, *Fgf1*, *Fos*, *Lamb3*, *Slc2a1*, *Sufu*, *Map2k2*, *Mtor*

Table 4.5 HF responsive KEGG disease pathway in the green module. The statistical results include: geometric mean of the individual p-values (p.geomean); mean of gene-set directional changes (stat.mean); global p-value (p.val); multiple test correction using Benjamini and Hochberg false discovery rate (q.val); the number of DEG (set.size). More details on statistics used in the columns are available in Luo et al. (2009).

4.3.5 Dominant response types

To associate either linear or nonlinear functions to all significantly over-represented biological processes, we considered as dominant the responses with the highest gene count per GOBP in each cluster (Table 4.4). Logarithmic and linear responses were the most dominant (Tables 4.4 and 4.6). Logarithmic response was characteristic of transport processes whereas linear response was mainly associated to metabolic processes.

Modules	Analysis	Regulation	Clusters					
			Transport Processes			Metabolic Processes		
			Response	Max Site	Min Site	Response	Max Site	Min Site
Green	HF Vs. LF	Up	Lin (1) Log (1)	2	1	Lin (3) Log (5)	10	1
		Down	Log (6)	5	1	Lin (39) Log (7)	1	1
		Both	Lin (3)	5	1	Lin (11) Log (2)	10	1
	LF Vs. Chow	Up	Log (4)	9	1	Lin (15) Log (11)	9	1
		Down	Log (1)	1, 9	2, 3, 7	Lin (2) Log (2)	1	10, 9
		Both	Lin (2)	9	3	Lin (8) Log (1)	9	9

Table 4.6 Dominant response types and intestinal sections. For each significant GOBP term, the response type with the highest gene count was defined as dominant. The Max site is the one where the highest number of differentially expressed genes associated to a specific GOBP term was found, while the Min site is the section where the lowest number of differentially expressed genes (for a given GOBP term) was observed.

4.3.6 Dominant intestinal sections

Within the first to the tenth section: site with the highest number of differentially expressed genes associated to a specific GOBP term was considered as actively responsive to diet,

which was called as Max site; and, the site with the lowest number of differentially expressed genes (for a given GOBP term) as being passively responsive to diet, which was termed as Min site (Table 4.6).With HF diet, sections 1-2, 5 and 10 were the most active. Sections 2 and 5 had the highest number of high expression values related to transport activities whereas sections 1 and 10 had the highest counts of high expression values specific to metabolic processes. With LF diet, sections 1-3, 7 and 9-10 were the most active. Sections 1 and 9 had the highest activity for both transport and metabolic processes.

4.4 Discussion

Over-consumption of energy dense fatty food combined with a sedentary and inactive lifestyle has evidenced a predisposition to chronic diseases (Cheung & Mao 2012; Desmarchelier et al. 2012; Walley et al. 2006). Studies have shown the importance of the gut in the maintenance of fat homeostasis and energy balance, with this regulation that is impaired in obesity and diabetes (Cheung & Mao 2012; Walley et al. 2006). The lipid sensing mechanism of small intestine is one of the first metabolic defenses against nutrient excess and contributes to maintain energy and metabolic balance (Wang et al. 2008). However, excessive fat intake is suggested to induce a delayed onset of satiety and promotes high fat diet hyperphagia (Little et al. 2007). The small intestine plays a key role in both absorption and digestion of dietary fat. Along the longitudinal axis, the small intestine shows regionalized differences in the absorptive capabilities and in the expression of genes encoding metabolic enzymes (Dusing et al. 2000; Gordon 1989). In consideration of the role of fatty acids as a metabolic fuel and as an integral part of the cell membrane (Gordon 1989; Niot et al. 2009), our main goals were: (1) to identify which metabolic and transport processes were fat-responsive; (2) to understand whether the prevalence of over-expressed metabolic and transport processes was unevenly distributed from the proximal to the distal small intestine. Our broadest objective was investigating which intestinal sections

displayed similar response patterns and whether they were commonly affected by the fat content. The regionalized behavior of the small intestine was modelled by proposing characteristic response patterns (either linear or nonlinear - i.e., logarithmic or quadratic) of diet sensitive biological processes to take into account how varying levels of fat could affect gene expression. Contrast statistical analysis between HF and LF showed significant impact of HF on the differential gene expression in small intestine. Irrespective of the response type, approximately 93% of the DEGs were HF responsive whereas about 6.9% responded to LF (Table 4.7). Taking into account the response of each DEG, a remarkable 43% was up-regulated and showed logarithmic response to HF (Table 4.7). This confirms high influence of HF within the intestinal milieu and suggests that excess fat in the lumen can be associated to a saturated response in terms of gene expression intensity. Excessive amounts of fat that reach the middle and the distal parts of small intestine can exceed the organ capacity. Such behavior describes a saturation response, which could set upper limits to digestion and absorption.

Test	Response	Regulation	DEG count	Relative %	Total %
HF vs. LF	Linear	Up	267	16.18	93.08
		Down	298	18.06	
	Quadratic	Up	36	2.18	
		Down	14	0.84	
	Logarithmic	Up	709	42.97	
		Down	212	12.85	
LF vs. Chow	Linear	Up	43	2.61	6.91
		Down	51	3.09	
	Quadratic	Up	8	0.48	
		Down	9	0.55	
	Logarithmic	Up	1	0.06	
		Down	2	0.12	

Table 4.7: Response specific DEG counts. Using the limma library we obtained diet and response specific DEGs along the longitudinal axis of murine small intestine (adjusted p-value = 0.1). We summarized the relative percentage of DEGs counts per response types.

The proximal small intestine is vital for nutrient digestion and absorption. During acute lipid intake, the upper intestinal lipid sensing mechanism suppresses glucose production (i.e., by the activation of intestine-brain axis) to negatively regulate food intake and maintain energy balance (Wang et al. 2008). Conversely, increased levels of carbohydrate intake suppress lipid metabolism but increase carbohydrate metabolism (Jeukendrup 2003). We observed an inverse relation between numerous biological processes that responded to either HF or LF along the longitudinal axis of the small intestine (Table 4.8). Both lipid metabolism and transport related processes were highly represented in HF while being repressed in LF. Processes related to carbohydrate metabolism and ion transport were up-regulated in LF and down-regulated in HF. The intestinal feedback varied considerably between HF and LF. In response to HF, we observed the prevalence of metabolic activities in the upper proximal and lower distal sections, whereas upper proximal and middle

sections had the highest number of transport activities. In LF diet, the upper proximal and lower distal sections were particularly active in both transport and metabolic processes. The proximal intestine was always particularly active, even with low levels of fat intake, while the most downstream sections were more sensitive to dietary fat. Linear and logarithmic response types were the most dominant in both metabolic and transport processes. The logarithmic response was more characteristic for the transport processes while the linear response was mainly associated to metabolic processes (Table 4.6).

BP	Rel	GOBP	HF					LF				
			pgmean	smean	p.val	q.val	ssize	pgmean	smean	p.val	q.val	ssize
MP	HFup vs. LFdown	GO:0044242 cellular lipid catabolic process	0.09	1.28	<0.01	0.01	12	0.08	-1.35	<0.01	0.01	12
		GO:0006644 phospholipid metabolic process	0.11	1.22	<0.01	0.01	14	0.10	-1.27	<0.01	0.01	14
		GO:0016042 lipid catabolic process	0.10	1.19	<0.01	0.01	14	0.10	-1.22	<0.01	0.01	14
	HFdown vs. LFup	GO:0010628 positive regulation of gene expression	0.11	-1.07	<0.01	0.01	27	0.23	0.66	0.02	0.09	27
		GO:0045944 positive regulation of transcription from RNA polymerase II promoter	0.12	-1.02	<0.01	0.01	22	0.21	0.71	0.01	0.07	22
		GO:0045893 positive regulation of transcription, DNA-dependent	0.11	-1.07	<0.01	0.01	27	0.23	0.66	0.02	0.09	27
		GO:0051254 positive regulation of RNA metabolic process	0.11	-1.07	<0.01	0.01	28	0.22	0.67	0.02	0.09	28
		GO:0051173 positive regulation of nitrogen compound metabolic process	0.11	-1.03	<0.01	0.01	34	0.23	0.67	0.02	0.09	34
		GO:0005996 monosaccharide metabolic process	0.11	-1.11	<0.01	<0.01	23	0.08	1.41	<0.01	<0.01	23
		GO:0019318 hexose metabolic process	0.11	-1.09	<0.01	0.01	22	0.09	1.35	<0.01	<0.01	22
		GO:0006006 glucose metabolic process	0.14	-0.99	<0.01	0.01	18	0.12	1.16	<0.01	<0.01	18
		GO:0044262 cellular carbohydrate metabolic process	0.24	-0.65	0.02	0.07	16	0.13	1.15	<0.01	<0.01	16
		GO:0010605 negative regulation of macromolecule metabolic process	0.10	-1.02	<0.01	0.01	36	0.20	0.80	0.01	0.04	36
		GO:0031324 negative regulation of cellular metabolic	0.12	-0.94	<0.01	0.01	39	0.19	0.83	<0.01	0.03	39

		process										
		GO:0005975 carbohydrate metabolic process	0.11	-1.00	<0.01	0.01	39	0.06	1.48	<0.01	<0.01	39
		GO:0016051 carbohydrate biosynthetic process	0.23	-0.71	0.01	0.05	10	0.09	1.38	<0.01	<0.01	10
		GO:0044723 single-organism carbohydrate metabolic process	0.13	-0.97	<0.01	0.01	30	0.08	1.40	<0.01	<0.01	30
	HFboth vs. LFboth	GO:0008202 steroid metabolic process	0.05	1.70	<0.01	<0.01	10	0.10	1.23	<0.01	0.01	10
		GO:0044255 cellular lipid metabolic process	0.07	1.42	<0.01	<0.01	45	0.22	0.75	0.01	0.09	45
		GO:0044283 small molecule biosynthetic process	0.07	1.40	<0.01	<0.01	30	0.13	0.96	<0.01	0.04	30
		GO:0006066 alcohol metabolic process	0.07	1.39	<0.01	<0.01	19	0.15	1.02	<0.01	0.02	19
		GO:0010565 regulation of cellular ketone metabolic process	0.10	1.31	<0.01	<0.01	11	0.13	1.17	<0.01	0.01	11
		GO:0042180 cellular ketone metabolic process	0.10	1.31	<0.01	<0.01	11	0.13	1.17	<0.01	0.01	11
		GO:1901615 organic hydroxy compound metabolic process	0.05	1.60	<0.01	<0.01	23	0.07	1.50	<0.01	<0.01	23
TP	Hfup vs. LFdown	GO:0016192 vesicle-mediated transport	0.04	1.69	<0.01	<0.01	35	0.06	-1.23	<0.01	0.01	35
	HFdown vs. LFup	GO:0055085 transmembrane transport	0.16	-0.83	<0.01	0.02	31	0.16	0.96	<0.01	0.01	31
		GO:0006812 cation transport	0.19	-0.83	<0.01	0.03	17	0.19	0.80	0.01	0.04	17
		GO:0006811 ion transport	0.18	-0.82	<0.01	0.03	30	0.19	0.76	0.01	0.06	30
		GO:0043269 regulation of ion transport	0.19	-0.79	0.01	0.03	10	0.22	0.68	0.02	0.09	10
	Hfboth vs. LFboth	GO:0071702 organic substance transport	0.14	0.95	<0.01	0.03	68	0.13	1.06	<0.01	0.02	68

Table 4.8 GOBP terms in HF and LF. Each row represents a particular biological process (BP), either metabolic (MP) or transport (TP). Often, such processes were characterized by an inverse relationship between HF and LF. The nature of each relation is in column "Rel". For example, the first row of column "Rel" represents three metabolic GOBP terms that were up-regulated in HF and down-regulated in LF. The statistical results include: geometric mean of the individual p-values (pgmean); mean of gene-set directional changes (smean); global p-value (p.val); multiple test correction using Benjamini and Hochberg false discovery rate (q.val); the number of DEGs (ssize). More details on statistics used in the columns are available in (Luo et al. 2009). The statistical results are reported for both HF (HF vs. LF) and LF (LF vs. Chow).

4.5 Conclusions

HF western style diet showed high impact on the differential gene expression in the small intestine of C57BL/6J mice. Increased HF intake repressed carbohydrate metabolism and led to over-expressing genes involved in lipid metabolism. The middle and the distal sections were particularly sensitive to fat content while the proximal section was proactive irrespective of the diet. The transport processes responded to increased levels of fat content with a saturated logarithmic response. In both HF and LF there was high metabolic activity in the upper proximal and in the lower distal small intestine although the enriched pathways showed opposite trends of regulation (i.e., lipid metabolic processes were up-regulated on a high-fat diet but the same processes were down-regulated on a LF diet). Such responses could be due to the lipid-mediated small intestine adaptation (i.e., increased levels of fat triggers higher absorption) and lead to chronic metabolic disorders.

Chapter 5

Conclusions

Nonlinear modelling can be used as an investigative tool to describe different types of physiological responses triggered by external (i.e., environmental) stimuli. This approach is particularly suitable to characterize the genetic mechanisms of expression involved in the regulation of biological processes as fat absorption and digestion in response to various dietary habits or to microbiota activity (i.e., with the term microbiota we indicate the ecological community of microorganisms of our body). In molecular nutrition, the small intestine plays a key role in the absorption and digestion of food, but its relevance has often been overlooked. The small-intestine consists of a large surface area that functions as the main portal for nutrient absorption, with physiological differences from the proximal to the distal section. The small intestine is colonized by a large and heterogeneous community of microorganisms (i.e., the microbiota that includes both beneficial and harmful populations) and performs the simultaneous functions of (1) nutrient absorption and (2) barrier against the entry of harmful antigens (such as pathogenic bacteria or allergens). Breaches in the small intestinal boundary disturb the intestinal homeostasis and excess high-fat intake compromises intestinal barrier sealing functions by down-regulating tight junction proteins (i.e., ZO-1 and occludin; see Cani et al. 2008). This triggers inflammatory damages until reaching extra-intestinal sites (e.g., the liver - Carding et al. 2015; Fasano &

Shea-Donohue 2005; see Figure 1.1). Such a complex interplay between physiological and environmental reactions (e.g., the relationships that link the diet to the activity of the luminal antigens) can be studied with nonlinear modelling. Also the modulation of gene expression mechanisms driven by fat availability can be studied effectively by using nonlinear models that mimic the consequences of fat overload (e.g., the saturation of dietary fat absorption in response to excessive amounts of fat intake can be described by modelling gene expression as a logarithmic function of different levels of fat intake).

5.1 Nonlinear regulation of fat intake

The importance of studying nonlinear responses to dietary fat intake in the small intestine is justified by evident morphological and physiological differences that characterize its longitudinal axis (DeSesso & Jacobson 2001). Also, the small intestine is capable of accommodating excess fat with increased absorption and this response is coordinated at the level of gene expression. Dietary fat mediated up-regulation of genes (i.e, FATP4, I-FABP or L-FABP) facilitates intestinal fatty acid uptake (Niot et al. 2009; Petit et al. 2007). This activity can be measured using nonlinear logarithmic or exponential functions that depict saturation or proactive absorption responses to fat intake, respectively. The approach based on nonlinear patterns perfectly matches the need of modelling the regionalized responses of small intestine to the intake of dietary fat gradients (i.e., by considering the modulations that involve absorptive and transport processes).

5.1.1 Consequences of fat intake on the transcriptomic response along the longitudinal axis of small intestine

The Chapter 3 of this thesis investigated the linear and nonlinear gene expression responses to dietary fat doses (i.e., 10%, 20%, 30% or 45% of kcal from fat) along the

longitudinal axis of the small intestine. The linear gene expression represents a constant raise in response to increasing fat intake. The logarithmic response illustrates the presence of overload mechanisms. The exponential response stands for enhanced gene expression behavior as a consequence of highest levels of fat intake. The quadratic response measures parabolic functions that are typical for enzyme kinetics. The cubic response refers to an oscillating behavior that can be associated to circadian rhythms. The study of Chapter 3 aimed at analyzing the presence of either gene-specific or section-specific responses to fat doses. A gene-specific response explains expression of genes exhibiting a constant pattern in each of the three intestinal sections. For example, we found that *Serpina1b* was up-regulated and linearly responsive to fat doses in the whole small intestine (Table 3.11). Section-specific responses characterize those genes that modulate their behavior depending on the intestine sections where the differential expression occurs. For example, *Abca1* changed from being linearly responsive in the proximal to be exponentially responsive in the middle section (always down-regulated). The main reasons that motivate the application of this novel approach are: (1) to provide dynamic understanding of the influence of dietary fat on intestinal lipid metabolism; (2) to understand which factors lead to fat mediated adaptations of intestinal lipid metabolism thus triggering the development of metabolic disorders (e.g., obesity).

We found that an increased intake of high-fat diets exerts significant impact on the differential gene expression patterns of the small intestine (Tables 3.2 and 4.7). The middle section is known to be the most responsive to fat intake (de Wit et al. 2011) and this was corroborated by the findings of our analysis. By studying both linear and nonlinear responsive genes, we showed that the middle section of the small intestine contained the highest number of differentially expressed genes. Conversely, the lowest number of differentially expressed genes was found in the distal section (Table 3.2). As illustrated by

the study of Chapter 4, the intake of a high-fat diet mainly resulted in the over-expression of processes related to metabolism and transport of lipid, while there was down-regulation of carbohydrate metabolism. Such pattern suggests that an adaptation mechanism is at play in the small intestine and can result in enhanced levels of fat absorption; this mechanism is likely to be driven by fat availability (Petit et al. 2007). The whole small intestine displayed the prevalence of linear gene expression responses as a function of increasing levels of dietary fat doses (Table 3.2 and Figure 3.3). The availability of highest amounts of fat in the proximal section stimulated the highest relative importance of saturated logarithmic responses (Table 3.2 and Figure 3.3). The overflow of (decreasing amounts of) excess fat into the distal region resulted in the highest relative importance of unsaturated, exponential responses (Table 3.2 and Figure 3.3). Along the intestinal axis, logarithmic and linear responses were the most representative in case of transport and metabolic processes, respectively (Table 4.6). In each of the three sections of the small intestine, there was gene-specific linear response of processes related to inflammatory reaction (Table 3.11). Processes related to cholesterol transport and efflux exhibited section-specific responses. In the middle section, there was inhibition (with negative, exponential response) of cholesterol transport and efflux processes in response to the intake of highest amounts of fat. The distal section showed an enhanced expression (with positive, exponential response) in terms of negative regulation of caspase activity (i.e., *Birc5*; see Table 3.11). This suggests that the small intestine activated some measures to counterbalance the excess lipid overflow into the distal section (Desmarchelier et al. 2012). In this way the intestine is able to cope by means of apoptosis the lipotoxic effect stimulated by excess fat (Desmarchelier et al. 2012). However, disease enrichment analysis suggested a significant over-representation of pathways related to cancer (Table 4.5).

5.2 Applicability beyond intestinal system

This thesis introduced the use of linear and nonlinear modelling to describe the changes of gene expression in response to dietary fat intake. Therefore, the novelties of the approach implemented in the Chapter 3 and 4 consist in: (1) describing the carrying capacity of each section of the small intestine in response to a gradient of dietary fat intake; (2) providing a dynamic characterization of the expression pattern of the genes, where each response pattern explains a unique state of gene expression to varying levels of fat intake; (3) illustrating the presence of gene- and section-specific responses to fat intake along the longitudinal axis of the small intestine. Thus, our study draws a map concerning the carrying capacity of genes and intestinal sections to different doses of dietary fat (i.e., using linear and nonlinear modelling, fat intake was studied as a continuous predictor of gene expression). As potential future applications, we suggest that this approach can be promising to investigate the differences in the efficacy of various treatments (e.g., a particular drug can be more effective in a group of individuals only and its consequences can be measured by the fact it determines unsaturated or saturated gene expression response patterns) or to model individual-level responses to increased levels of nutrients (e.g., in the case of fat intake it can be used to investigate for which reasons some individuals are more resistant to weight gain than others). These studies (i.e., based on the differences among individuals) could serve to identify peculiar response types in specific genes as potential biomarkers. Such biomarkers could model continuous responses (i.e., gene expression) as a function of external, environmental factors (e.g., various drug doses or different concentrations of micronutrients) and serve to examine: (1) differences in response patterns among individuals; (2) spatial or temporal trends in the responses within different regions of the organs (e.g., in this thesis we have illustrated that the small intestine displays regionalized functions).

5.3 Outlook

The investigation of the diet-mediated physiological and pathological changes represents the key for better understanding the onset of metabolic disorders. Previous studies have demonstrated high-fat mediated long term effects such as chronic inflammation, weight gain and obese phenotype (Cani et al. 2007; Garrett et al. 2010; de Wit et al. 2011), showing the high individual-level variability that characterizes these responses (e.g., individualized response to weight gain; see Speakman et al. 2011). In my work, the transcriptomic response was modelled as a function of dietary fat doses, and DEGs displayed regionalized expression along the longitudinal axis of small intestine. A possible extension of the study could be based on (1) investigating individualized responses to treatments, and (2) identifying specific biomarkers that are robust to individual-level variability. Also epigenetics could be taken into account to consider how high-fat diets influence the methylation state of genes (Amaral et al. 2014; Milagro et al. 2009).
The large surface area of intestine is a hotbed for both beneficial and harmful pathogens (O'Hara & Shanahan 2006; Shen et al. 2013). Influence of diet on the composition of gut microbiota has been widely documented, with western-style high-fat diets that reduce bacterial diversity and lead to dysbiosis (Bäckhed et al. 2004; Bäckhed et al. 2007). This triggers the development of metabolic complications such as obesity, inflammatory bowel disease and diabetes (Clemente et al. 2012; Shen et al. 2013), suggesting that gene expression can be influenced by various external factors (e.g., diet, microbiome and their interactions). Thus, gene expression in intestine depends on: (1) cell types (i.e., enteroendocrine vs. absorptive cells), (2) location (i.e., the crypt-villus axis), and (3) luminal factors (e.g., diet and microbiota). This thesis takes into account the role of diet on gene expression and could be integrated by new, multi-factorial experiments to consider the interactions between diet and microbiota.

Bibliography

Abreu, M.T., 2010. Toll-like receptor signalling in the intestinal epithelium: how bacterial recognition shapes intestinal function. *Nature Reviews Immunology*, 10(2), pp.131–44.

Al-Dwairi, A. et al., 2014. Enhanced gastrointestinal expression of cytosolic malic enzyme (ME1) induces intestinal and liver lipogenic gene expression and intestinal cell proliferation in mice. *PloS One*, 9(11), p.e113058.

Altmann, S.W. et al., 2004. Niemann-Pick C1 Like 1 protein is critical for intestinal cholesterol absorption. *Science*, 303(5661), pp.1201–4.

Amaral, C.L. et al., 2014. DNA Methylation Changes Induced by a High-Fat Diet and Fish Oil Supplementation in the Skeletal Muscle of Mice. *Journal of Nutrigenetics and Nutrigenomics*, 7(4-6), pp.314–26.

van Ampting, M.T.J. et al., 2012. Intestinally secreted C-type lectin Reg3b attenuates salmonellosis but not listeriosis in mice. *Infection and Immunity*, 80(3), pp.1115–20.

Bäckhed, F. et al., 2007. Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. *Proceedings of the National Academy of Sciences*, 104(3), pp.979–84.

Bäckhed, F. et al., 2004. The gut microbiota as an environmental factor that regulates fat storage. *Proceedings of the National Academy of Sciences of the United States of America*, 101(44), pp.15718–23.

Balwierz, A. et al., 2009. Angiogenesis in the New Zealand obese mouse model fed with high fat diet. *Lipids in Health and Disease*, 8(1), p.13.

Balzan, S. et al., 2007. Bacterial translocation: overview of mechanisms and clinical impact. *Journal of Gastroenterology and Hepatology*, 22(4), pp.464–71.

Bellafante, E. et al., 2014. PGC-1β promotes enterocyte lifespan and tumorigenesis in the intestine. *Proceedings of the National Academy of Sciences of the United States of America*, 111(42), pp.E4523–31.

Benjamini, Y. & Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*. *Series*

B (Methodological), 57(1), pp.289–300.

Berge, K.E. et al., 2000. Accumulation of dietary cholesterol in sitosterolemia caused by mutations in adjacent ABC transporters. *Science*, 290(5497), pp.1771–5.

Blanc-Brude, O.P. et al., 2007. IAP survivin regulates atherosclerotic macrophage survival. *Arteriosclerosis, thrombosis, and vascular biology*, 27(4), pp.901–7.

Bloks, V.W. et al., 2004. Down-regulation of hepatic and intestinal Abcg5 and Abcg8 expression associated with altered sterol fluxes in rats with streptozotocin-induced diabetes. *Diabetologia*, 47(1), pp.104–12.

Bolstad, B., 2001. Probe level quantile normalization of high density oligonucleotide array data. *Unpublished manuscript*.

Bolstad, B.M. et al., 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, 19(2), pp.185–93.

Bolstad, B.M., 2004. *Low-level Analysis of High-density Oligonucleotide Array Data: Background, Normalization and Summarization*. UNIVERSITY OF CALIFORNIA, BERKELEY.

Borgström, B., 1977. Digestion and absorption of lipids. *International Review of Physiology*, 12, pp.305–23.

Bourgon, R., Gentleman, R. & Huber, W., 2010. Independent filtering increases detection power for high-throughput experiments. *Proceedings of the National Academy of Sciences of the United States of America*, 107(21), pp.9546–51.

Brettschneider, J. et al., 2008. Quality assessment for short oligonucleotide microarray data. *Technometrics*, 50(3).

Brun, P. et al., 2007. Increased intestinal permeability in obese mice: new evidence in the pathogenesis of nonalcoholic steatohepatitis. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 292(2), pp.G518–25.

Buhman, K.K. et al., 2000. Resistance to diet-induced hypercholesterolemia and gallstone formation in ACAT2-deficient mice. *Nature Medicine*, 6(12), pp.1341–7.

Buts, J.P. et al., 1990. Refeeding after starvation in the rat: comparative effects of lipids, proteins and carbohydrates on jejunal and ileal mucosal adaptation. *European Journal of Clinical Investigation*, 20(4), pp.441–52.

Cani, P.D. et al., 2008. Changes in gut microbiota control metabolic endotoxemia-induced inflammation in high-fat diet-induced obesity and diabetes in mice. *Diabetes*, 57(6), pp.1470–81.

Cani, P.D. et al., 2007. Metabolic endotoxemia initiates obesity and insulin resistance. *Diabetes*, 56(7), pp.1761–72.

Carding, S. et al., 2015. Dysbiosis of the gut microbiota in disease. *Microbial Ecology in Health and Disease*, 26, p.26191.

Chen, M. et al., 2001. Gut expression and regulation of FAT/CD36: possible role in fatty acid transport in rat enterocytes. *American Journal of Physiology-Endocrinology and Metabolism*, 281(5), pp.E916–23.

Cheung, W.W. & Mao, P., 2012. Recent advances in obesity: genetics and beyond. *ISRN Endocrinology*, 2012, p.536905.

Chiou, S.-K. et al., 2003. Survivin expression in the stomach: implications for mucosal integrity and protection. *Biochemical and Biophysical Research Communications*, 305(2), pp.374–379.

Chu, F.-F. et al., 2004. Bacteria-induced intestinal cancer in mice with disrupted Gpx1 and Gpx2 genes. *Cancer Research*, 64(3), pp.962–8.

Clark, S.B. & Tercyak, A.M., 1984. Reduced cholesterol transmucosal transport in rats with inhibited mucosal acyl CoA:cholesterol acyltransferase and normal pancreatic function. *Journal of Lipid Research*, 25(2), pp.148–59.

Clemente, J.C. et al., 2012. The impact of the gut microbiota on human health: An integrative view. *Cell*, 148(6), pp.1258–1270.

Csardi, G. & Nepusz, T., 2006. The igraph software package for complex network research. *InterJournal, Complex Systems*, 1695(5), pp.1–9.

Daves, M.H. et al., 2011. Meta-analysis of multiple microarray datasets reveals a common gene signature of metastasis in solid tumors. *BMC Medical Genomics*, 4(1), p.56.

Dawson, K., Rodriguez, R.L. & Malyj, W., 2005. Sample phenotype clusters in high-density oligonucleotide microarray data sets are revealed using Isomap, a nonlinear algorithm. *BMC Bioinformatics*, 6, p.195.

Delzenne, N.M. & Cani, P.D., 2011. Interaction between obesity and the gut microbiota: relevance in nutrition. *Annual Review of Nutrition*, 31, pp.15–31.

DeSesso, J.M. & Jacobson, C.F., 2001. Anatomical and physiological parameters affecting gastrointestinal absorption in humans and rats. *Food and Chemical Toxicology*, 39(3), pp.209–28.

Desmarchelier, C. et al., 2012. C57Bl/6 N mice on a western diet display reduced intestinal and hepatic cholesterol levels despite a plasma hypercholesterolemia. *BMC Genomics*, 13(1), p.84.

Desvergne, B., Michalik, L. & Wahli, W., 2006. Transcriptional regulation of metabolism. *Physiological Reviews*, 86(2), pp.465–514.

Ding, S. et al., 2010. High-fat diet: bacteria interactions promote intestinal inflammation which precedes and correlates with obesity and insulin resistance in mouse. *PloS One*,

5(8), p.e12191.

Ding, S. & Lund, P.K., 2011. Role of intestinal inflammation as an early event in obesity and insulin resistance. *Current Opinion in Clinical Nutrition and Metabolic Care*, 14(4), pp.328–33.

Dobbs, R. et al., 2014. *Overcoming obesity : An initial economic analysis*,

Drover, V.A. et al., 2005. CD36 deficiency impairs intestinal lipid secretion and clearance of chylomicrons from the blood. *Journal of Clinical Investigation*, 115(5), pp.1290–7.

Dunel-Erb, S. et al., 2001. Restoration of the jejunal mucosa in rats refed after prolonged fasting. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 129(4), pp.933–47.

Duplus, E., Glorian, M. & Forest, C., 2000. Fatty acid regulation of gene transcription. *Journal of Biological Chemistry*, 275(40), pp.30749–52.

Dusing, M.R. et al., 2000. A duodenum-specific enhancer regulates expression along three axes in the small intestine. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 279(5), pp.G1080–93.

Eckburg, P.B. et al., 2005. Diversity of the human intestinal microbial flora. *Science*, 308(5728), pp.1635–8.

Everard, A. et al., 2013. Cross-talk between Akkermansia muciniphila and intestinal epithelium controls diet-induced obesity. *Proceedings of the National Academy of Sciences*, 110(22), pp.9066–71.

Everard, A. et al., 2011. Responses of gut microbiota and glucose and lipid metabolism to prebiotics in genetic obese and diet-induced leptin-resistant mice. *Diabetes*, 60(11), pp.2775–86.

Fasano, A. et al., 2000. Zonulin, a newly discovered modulator of intestinal permeability, and its expression in coeliac disease. *The Lancet*, 355(9214), pp.1518–9.

Fasano, A. & Shea-Donohue, T., 2005. Mechanisms of disease: the role of intestinal barrier function in the pathogenesis of gastrointestinal autoimmune diseases. *Nature Clinical Practice Gastroenterology & Hepatology*, 2(9), pp.416–22.

De Filippo, C. et al., 2010. Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proceedings of the National Academy of Sciences*, 107(33), pp.14691–6.

Fleissner, C.K. et al., 2010. Absence of intestinal microbiota does not protect mice from diet-induced obesity. *British Journal of Nutrition*, 104(6), pp.919–29.

Fröhlich, H. et al., 2007. GOSim--an R-package for computation of information theoretic GO similarities between terms and gene products. *BMC Bioinformatics*, 8, p.166.

Fukuda, S. & Pelus, L.M., 2006. Survivin, a cancer target with an emerging role in normal

adult tissues. *Molecular Cancer Therapeutics*, 5(5), pp.1087–98.

Garrett, W.S., Gordon, J.I. & Glimcher, L.H., 2010. Homeostasis and Inflammation in the Intestine. *Cell*, 140(6), pp.859–870.

Gentleman, R. et al., 2012. genefilter: Methods for filtering genes from microarray experiments. *R package version 1.36.0*.

Gentleman, R.C. et al., 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology*, 5(10), p.R80.

George, M.D. et al., 2008. In vivo gene expression profiling of human intestinal epithelial cells: analysis by laser microdissection of formalin fixed tissues. *BMC Genomics*, 9, p.209.

Georgiadi, A. & Kersten, S., 2012. Mechanisms of gene regulation by fatty acids. *Advances in Nutrition: An International Review Journal*, 3(2), pp.127–34.

Gniuli, D. et al., 2008. High saturated-fat diet induces apoptosis in rat enterocytes and blunts GIP and insulin-secretive response to oral glucose load. *International Journal of Obesity*, 32(5), pp.871–4.

Goncalves, A. et al., 2015. Fat-soluble vitamin intestinal absorption: absorption sites in the intestine and interactions for absorption. *Food Chemistry*, 172, pp.155–60.

Gordon, J.I., 1989. Intestinal epithelial differentiation: new insights from chimeric and transgenic mice. *The Journal of Cell Biology*, 108(4), pp.1187–94.

Greenwalt, D.E., Scheck, S.H. & Rhinehart-Jones, T., 1995. Heart CD36 expression is increased in murine models of diabetes and in mice fed a high fat diet. *Journal of Clinical Investigation*, 96(3), pp.1382–8.

Harris, L.W. et al., 2009. Gene expression in the prefrontal cortex during adolescence: implications for the onset of schizophrenia. *BMC Medical Genomics*, 2, p.28.

Hartmann, P. et al., 2013. Deficiency of intestinal mucin-2 ameliorates experimental alcoholic liver disease in mice. *Hepatology*, 58(1), pp.108–19.

Herrmann, T. et al., 2001. Mouse fatty acid transport protein 4 (FATP4): characterization of the gene and functional assessment as a very long chain acyl-CoA synthetase. *Gene*, 270(1-2), pp.31–40.

Hollander, D. & Truscott, T.C., 1976. Mechanism and site of small intestinal uptake of vitamin D3 in pharmacological concentrations. *The American Journal of Clinical Nutrition*, 29(9), pp.970–5.

Hotamisligil, G.S., 2006. Inflammation and metabolic disorders. *Nature*, 444(7121), pp.860–7.

Institute of Medicine, 2005. *Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids (Macronutrients)*,

Washington, D.C.: National Academies Press.

Iqbal, J. & Hussain, M.M., 2009. Intestinal lipid absorption. *American Journal of Physiology-Endocrinology and Metabolism*, 296(6), pp.E1183–94.

Irizarry, R. a. et al., 2003. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Research*, 31(4), p.15e–15.

Ivanov, I.I. et al., 2009. Induction of intestinal Th17 cells by segmented filamentous bacteria. *Cell*, 139(3), pp.485–98.

Janssen, S. & Depoortere, I., 2013. Nutrient sensing in the gut: new roads to therapeutics? *Trends in Endocrinology and Metabolism*, 24(2), pp.92–100.

Jeukendrup, A.E., 2003. Modulation of carbohydrate and fat utilization by diet, exercise and environment. *Biochemical Society Transactions*, 31(Pt 6), pp.1270–3.

Kaminsky, L.S. & Zhang, Q.-Y., 2003. The small intestine as a xenobiotic-metabolizing organ. *Drug Metabolism and Disposition*, 31(12), pp.1520–5.

Kaufman, L. & Rousseeuw, P.J., 1990. *Finding Groups in Data: An Introduction to Cluster Analysis*, New York: John Wiley & Sons, Inc.

Khaodhiar, L., McCowen, K.C. & Blackburn, G.L., 1999. Obesity and its comorbid conditions. *Clinical Cornerstone*, 2(3), pp.17–31.

Khovidhunkit, W. et al., 2003. Endotoxin down-regulates ABCG5 and ABCG8 in mouse liver and ABCA1 and ABCG1 in J774 murine macrophages: differential role of LXR. *Journal of Lipid Research*, 44(9), pp.1728–36.

Khovidhunkit, W. et al., 2000. Infection and inflammation-induced proatherogenic changes of lipoproteins. *The Journal of Infectious Diseases*, 181 Suppl, pp.S462–72.

Knights, D., Lassen, K.G. & Xavier, R.J., 2013. Advances in inflammatory bowel disease pathogenesis: linking host genetics and the microbiome. *Gut*, 62(10), pp.1505–10.

Kondo, H. et al., 2006. Differential regulation of intestinal lipid metabolism-related genes in obesity-resistant A/J vs. obesity-prone C57BL/6J mice. *American Journal of Physiology-Endocrinology And Metabolism*, 291(5), pp.E1092–9.

Kroll, K.M., Barkema, G.T. & Carlon, E., 2008. Modeling background intensity in DNA microarrays. *arXiv preprint arXiv:0712.3494*, 77(6), p.061915.

de La Serre, C.B. et al., 2010. Propensity to high-fat diet-induced obesity in rats is associated with changes in the gut microbiota and gut inflammation. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 299(2), pp.G440–8.

Langfelder, P. & Horvath, S., 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, 9, p.559.

de Lartigue, G., de La Serre, C.B. & Raybould, H.E., 2011. Vagal afferent neurons in high

fat diet-induced obesity; intestinal microflora, gut inflammation and cholecystokinin. *Physiology & Behavior*, 105(1), pp.100–5.

Ley, R.E. et al., 2006. Microbial ecology: human gut microbes associated with obesity. *Nature*, 444(7122), pp.1022–3.

Lim, S.S. et al., 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet*, 380(9859), pp.2224–60.

Lipshutz, R.J. et al., 1999. High density synthetic oligonucleotide arrays. *Nature Genetics*, 21(1 Suppl), pp.20–4.

Little, T.J., Horowitz, M. & Feinle-Bisset, C., 2007. Modulation by high-fat diets of gastrointestinal function and hormones associated with the regulation of energy intake: implications for the pathophysiology of obesity. *The American Journal of Clinical Nutrition*, 86(3), pp.531–41.

Lombardo, D., Fauvel, J. & Guy, O., 1980. Studies on the substrate specificity of a carboxyl ester hydrolase from human pancreatic juice. I. Action on carboxyl esters, glycerides and phospholipids. *Biochimica et Biophysica Acta (BBA)-Enzymology*, 611(1), pp.136–46.

Lozupone, C. a et al., 2012. Diversity, stability and resilience of the human gut microbiota. *Nature*, 489(7415), pp.220–30.

Luo, W. et al., 2009. GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics*, 10, p.161.

Mansbach, C.M. & Dowell, R., 2000. Effect of increasing lipid loads on the ability of the endoplasmic reticulum to transport lipid to the Golgi. *Journal of Lipid Research*, 41(4), pp.605–12.

Mansbach, C.M. & Nevin, P., 1998. Intracellular movement of triacylglycerols in the intestine. *Journal of Lipid Research*, 39(5), pp.963–8.

Martin, F.-P.J. et al., 2007. A top-down systems biology view of microbiome-mammalian metabolic interactions in a mouse model. *Molecular Systems Biology*, 3(112), p.112.

McCall, M.N., Bolstad, B.M. & Irizarry, R.A., 2010. Frozen robust multiarray analysis (fRMA). *Biostatistics*, 11(2), pp.242–53.

Meresse, B., Malamut, G. & Cerf-Bensussan, N., 2012. Celiac disease: an immunological jigsaw. *Immunity*, 36(6), pp.907–19.

Milagro, F.I. et al., 2009. High fat diet-induced obesity modifies the methylation pattern of leptin promoter in rats. *Journal of Physiology and Biochemistry*, 65(1), pp.1–9.

Mowat, A.M., 2003. Anatomical basis of tolerance and immunity to intestinal antigens.

Nature Reviews Immunology, 3(4), pp.331–41.

Mowat, A.M. & Agace, W.W., 2014. Regional specialization within the intestinal immune system. *Nature Reviews Immunology*, 14(10), pp.667–85.

Murphy, D., 2002. Gene expression studies using microarrays: principles, problems, and prospects. *Advances in Physiology Education*, 26(1-4), pp.256–70.

Nassir, F. et al., 2007. CD36 is important for fatty acid and cholesterol uptake by the proximal but not distal intestine. *Journal of Bological Chemistry*, 282(27), pp.19493–501.

Nassir, F. & Abumrad, N., 2009. CD36 and Intestinal Fatty Acid Absorption. *Immunology, Endocrine & Metabolic Agents in Medicinal Chemistry*, 9(1), pp.3–10.

Newman, L., Haryono, R. & Keast, R., 2013. Functionality of fatty acid chemoreception: a potential factor in the development of obesity? *Nutrients*, 5(4), pp.1287–300.

Ng, M. et al., 2014. Global, regional, and national prevalence of overweight and obesity in children and adults during 1980-2013: a systematic analysis for the Global Burden of Disease Study 2013. *The Lancet*, 384.

Nicholson, J.K. et al., 2004. The challenges of modeling mammalian biocomplexity. *Nature Biotechnology*, 22(10), pp.1268–74.

Niot, I. et al., 2009. Intestinal absorption of long-chain fatty acids: evidence and uncertainties. *Progress in Lipid Research*, 48(2), pp.101–15.

O'Hara, A.M. & Shanahan, F., 2006. The gut flora as a forgotten organ. *EMBO reports*, 7(7), pp.688–93.

Ockner, R.K. et al., 1972. A binding protein for fatty acids in cytosol of intestinal mucosa, liver, myocardium, and other tissues. *Science*, 177(4043), pp.56–8.

Ockner, R.K. & Manning, J.A., 1974. Fatty acid-binding protein in small intestine. Identification, isolation, and evidence for its role in cellular fatty acid transport. *Journal of Clinical Investigation*, 54(2), pp.326–38.

Pácha, J., 2000. Development of intestinal transport function in mammals. *Physiological Reviews*, 80(4), pp.1633–67.

Parks, B.W. et al., 2013. Genetic control of obesity and gut microbiota composition in response to high-fat, high-sucrose diet in mice. *Cell Metabolism*, 17(1), pp.141–152.

Peterson, L.W. & Artis, D., 2014. Intestinal epithelial cells: regulators of barrier function and immune homeostasis. *Nature Reviews Immunology*, 14(3), pp.141–53.

Petit, V. et al., 2007. Chronic high-fat diet affects intestinal fat absorption and postprandial triglyceride levels in the mouse. *Journal of Lipid Research*, 48(2), pp.278–87.

Phan, C.T. & Tso, P., 2001. Intestinal lipid absorption and transport. *Frontiers in*

Bioscience: a Journal and Virtual Library, 6, pp.D299–319.

Piccolo, S.R. et al., 2013. A single-sample microarray normalization method to facilitate personalized-medicine workflows. *Genomics*, 100(6), pp.337–344.

Pocai, A. et al., 2006. Restoration of hypothalamic lipid sensing normalizes energy and glucose homeostasis in overfed rats. *Journal of Clinical Investigation*, 116(4), pp.1081–91.

Poirier, H. et al., 1996. Localization and Regulation of the Putative Membrane Fatty-Acid Transporter (FAT) in the Small Intestine. *European Journal of Biochemistry*, 238(2), pp.368–373.

Qin, J. et al., 2010. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*, 464(7285), pp.59–65.

R Core Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Raymond, F. et al., 2010. Comparative gene expression profiling between human cultured myotubes and skeletal muscle tissue. *BMC Genomics*, 11, p.125.

Schnabl, B., 2013. Linking intestinal homeostasis and liver disease. *Current Opinion in Gastroenterology*, 29(3), pp.264–70.

Schulthess, G. et al., 1994. Absorption of monoacylglycerols by small intestinal brush border membrane. *Biochemistry*, 33(15), pp.4500–8.

Schwenk, R.W. et al., 2010. Fatty acid transport across the cell membrane: regulation by fatty acid transporters. *Prostaglandins, Leukotrienes and Essential Fatty Acids (PLEFA)*, 82(4-6), pp.149–54.

Serino, M. et al., 2009. Intestinal microflora and metabolic diseases. *Diabetes & Metabolism*, 35(4), pp.262–72.

Shaw-Smith, C.J. & Walters, J.R., 1997. Regional expression of intestinal genes for nutrient absorption. *Gut*, 40(1), pp.5–8.

Shen, J., Obin, M.S. & Zhao, L., 2013. The gut microbiota, obesity and insulin resistance. *Molecular Aspects of Medicine*, 34(1), pp.39–58.

Shi, H. et al., 2006. TLR4 links innate immunity and fatty acid-induced insulin resistance. *Journal of Clinical Investigation*, 116(11), pp.3015–25.

Shim, J. et al., 2009. Fatty acid transport protein 4 is dispensable for intestinal lipid absorption in mice. *Journal of Lipid Research*, 50(3), pp.491–500.

Simon, T. et al., 2011. Impact of murine intestinal apolipoprotein A-IV expression on regional lipid absorption, gene expression, and growth. *Journal of Lipid Research*, 52(11), pp.1984–94.

Smyth, G.K., 2005. Limma : Linear Models for Microarray Data. In R. Gentleman et al., eds. *Bioinformatics and Computational Biology Solutions using R and Bioconductor*. Springer New York, pp. 397–420.

Speakman, J.R. et al., 2011. Set points, settling points and some alternative models: theoretical options to understand how genes and environments combine to regulate body adiposity. *Disease Models & Mechanisms*, 4(6), pp.733–45.

Stahl, A. et al., 1999. Identification of the major intestinal fatty acid transport protein. *Molecular Cell*, 4(3), pp.299–308.

Stark, C. et al., 2011. The BioGRID Interaction Database: 2011 update. *Nucleic Acids Research*, 39(Database issue), pp.D698–704.

Stenman, L.K., Holma, R. & Korpela, R., 2012. High-fat-induced intestinal permeability dysfunction associated with altered fecal bile acids. *World Journal of Gastroenterology: WJG*, 18(9), pp.923–9.

Stevens, A. et al., 2013. Human growth is associated with distinct patterns of gene expression in evolutionarily conserved networks. *BMC Genomics*, 14(1), p.547.

Suzuki, T. & Hara, H., 2010. Dietary fat and bile juice, but not obesity, are responsible for the increase in small intestinal permeability induced through the suppression of tight junction protein expression in LETO and OLETF rats. *Nutrition & Metabolism*, 7, p.19.

Thomas, K.E. et al., 2006. Gliadin Stimulation of Murine Macrophage Inflammatory Gene Expression and Intestinal Permeability Are MyD88-Dependent: Role of the Innate Immune Response in Celiac Disease. *The Journal of Immunology*, 176(4), pp.2512–2521.

Thomson, A.B. et al., 1986. Dietary fat selectively alters transport properties of rat jejunum. *Journal of Clinical Investigation*, 77(1), pp.279–88.

Turnbaugh, P.J. et al., 2009. A core gut microbiome in obese and lean twins. *Nature*, 457(7228), pp.480–484.

Turnbaugh, P.J. et al., 2008. Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome. *Cell Host & Microbe*, 3(4), pp.213–23.

Turner, J.R., 2009. Intestinal mucosal barrier function in health and disease. *Nature Reviews Immunology*, 9(11), pp.799–809.

Vaishnava, S. et al., 2008. Paneth cells directly sense gut commensals and maintain homeostasis at the intestinal host-microbial interface. *Proceedings of the National Academy of Sciences*, 105(52), pp.20858–63.

Vassileva, G. et al., 2000. The intestinal fatty acid binding protein is not essential for dietary fat absorption in mice. *The FASEB Journal*, 14(13), pp.2040–6.

Velagapudi, V.R. et al., 2010. The gut microbiota modulates host energy and lipid metabolism in mice. *Journal of Lipid Research*, 51(5), pp.1101–1112.

Walley, A.J., Blakemore, A.I.F. & Froguel, P., 2006. Genetics of obesity and the prediction of risk for health. *Human Molecular Genetics*, 15(2), pp.R124–30.

Wang, P.Y.T. et al., 2008. Upper intestinal lipids trigger a gut-brain-liver axis to regulate glucose production. *Nature*, 452(7190), pp.1012–6.

Wang, T.Y. et al., 2013. New insights into the molecular mechanism of intestinal fatty acid absorption. *European Journal of Clinical Investigation*, 43(11), pp.1203–23.

Wang, X. et al., 2011. HTSanalyzeR: an R/Bioconductor package for integrated network analysis of high-throughput screens. *Bioinformatics*, 27(6), pp.879–80.

Werner, A., Kuipers, F. & Verkade, H.J., 2000. Fat Absorption and Lipid Metabolism in Cholestasis. In *Madame Curie Bioscience Database [Internet]*. Austin (TX): Landes Bioscience.

Westergaard, H. & Dietschy, J.M., 1974. Delineation of the dimensions and permeability characteristics of the two major diffusion barriers to passive mucosal uptake in the rabbit intestine. *Journal of Clinical Investigation*, 54(3), p.718.

Westergaard, H. & Dietschy, J.M., 1976. The mechanism whereby bile acid micelles increase the rate of fatty acid and cholesterol uptake into the intestinal mucosal cell. *Journal of Clinical Investigation*, 58(1), pp.97–108.

de Wit, N.J. et al., 2008. The role of the small intestine in the development of dietary fat-induced obesity and insulin resistance in C57BL/6J mice. *BMC Medical Genomics*, 1(1), p.14.

de Wit, N.J.W. et al., 2011. Dose-dependent effects of dietary fat on development of obesity in relation to intestinal differential gene expression in C57BL/6J mice. *PLoS One*, 6(4), p.e19145.

Wright, S.D. et al., 1990. CD14, a receptor for complexes of lipopolysaccharide (LPS) and LPS binding protein. *Science*, 249(4975), pp.1431–3.

Wu, J. et al., 2012. gcrma: Background Adjustment Using Sequence Information. *R package version 2.26.0*.

Wu, J., Irizarry, R. & Gentry, J.M.J., 2007. gcrma: Background Adjustment Using Sequence Information. *R package version*.

Yatsunenko, T. et al., 2012. Human gut microbiome viewed across age and geography. *Nature*, 486(7402), pp.222–7.

Yu, L. et al., 2002. Disruption of Abcg5 and Abcg8 in mice reveals their crucial role in biliary cholesterol secretion. *Proceedings of the National Academy of Sciences*, 99(25), pp.16237–42.

Zhang, H. et al., 2009. Human gut microbiota in obesity and after gastric bypass. *Proceedings of the National Academy of Sciences*, 106(7), pp.2365–70.

Zhong, Y. et al., 2015. Barley malt increases hindgut and portal butyric acid, modulates gene expression of gut tight junction proteins and Toll-like receptors in rats fed high-fat diets, but high advanced glycation end-products partially attenuate the effects. *Food & Function*, 6(9), pp.3165–76.

Zoetendal, E.G., Akkermans, A.D. & De Vos, W.M., 1998. Temperature gradient gel electrophoresis analysis of 16S rRNA from human fecal samples reveals stable and host-specific communities of active bacteria. *Applied and Environmental Microbiology*, 64(10), pp.3854–9.