Activation mechanism of AhR pathway in intestinal epithelial cells by commensal bacteria

Ludovica Marinelli

Sorbonne Université

Ecole doctorale 394 : Physiologie, Physiopathologie et Therapeutique

INRA Jouy-en-Josas, Institut MICALIS, Equipe FInE

Activation mechanism of AhR pathway in intestinal epithelial cells by commensal bacteria

Par Marinelli Ludovica

Thèse de Doctorat en Biologie

Dirigée par le Dr Blottière M. Hervé et le Dr Lapaque Nicolas

Présentée et soutenue publiquement le 30 mai 2018

Devant un jury composé de :

CLEMENT Karine,	Professeur des Universités, Paris 6	Président du Jury
SEGAIN Jean-Pierre,	Directeur de Recherche, INRA	Rapporteur
THEODOROU Vassilia,	Professeur des Universités, Toulouse	Rapporteur
COUMOUL Xavier,	Professeur des Universités, Paris 5	Examinateur
SOKOL Harry,	Professeur des Universités, Paris 6	Examinateur
BLOTTIERE Hervé,	Directeur de Recherche, INRA	Directeur de Thèse
LAPAQUE Nicolas,	Chargé de Recherche, INRA	Directeur de Thèse

Ai miei gentitori,

che mi hanno insegnato ad affrontare ogni sfida con il sorriso

Acknowledgments

Je remercie chacun des membres du jury de me faire l'honneur d'évaluer ce travail:

- Mme le Professeur Karine Clement, merci de présider ce jury et de partager votre expertise ;
- M le Docteur Jean-Pierre Segain et Mme le Professeur Vassilia Theodorou, merci d'avoir accepté d'etre rapporteur de ma thèse et d'apporter vos corrections afin d'améliorer ce travail.
- M le Professeur Xavier Coumoul et M le Professeur Harry Sokol, merci d'avoir accepté d'examiner ce travaille et de partager vos competences sur AhR.
- M le Docteur Hervé Blottière et M le Docteur Nicolas Lapaque, mes directeurs de thèse. Merci à Hervé de m'avoir accepté en thèse et de m'avoir proposé ce sujet. Merci a Nico de me avoir encadré a la paillasse comme au coin café, de m'avoir donné pleines des conseilles, de correctionnes et de m'avoir aussi stimulé avec ton interet scientifique. Merci a tous les deux de m'avoir fait confiance et de m'avoir permis de completer ce beau parcours de thèse.

Table of Contents

Table of Figures

16973

Chapter 1. Introduction

1.1. *The gastrointestinal tract*

The gastrointestinal tract (GI) has extensively been studied for its digestive functions. Together with accessory digestive organs, the GI tract constitutes the digestive system, whose main function consists in digestion, absorption and subsequent distribution of nutrients through the body. As a result, the GI tract is divided into functional regions, each of which is characterized by its own physico-chemical conditions (*Figure 1*).

The ingested food is mechanically broken down in the oral cavity and mixed to the first digestive enzymes (amylases and lipases). Through the oesophagus, the food reaches the stomach where other digestive enzymes (pepsins) are released along with hydrochloride acid, causing the pH to drop (pH 1-2). The food content passes then through the different sections of the small intestine (duodenum, jejunum and ileum) where the pH increases as bicarbonate is releases (pH5.7-7.3). During this passage other digestive enzymes are released (pancreatic enzymes, mucosal enzymes and bile salts), resulting in the cleavage of carbohydrates, proteins, and lipids and the absorption of low molecular amino acids, simple sugars and fatty acids. From the ileum, non-degradable food components land in the large intestine, traversing the cecum and the colon (ascending, transverse and descending; pH 5.7-6.8) where water and salts are absorbed and the fermentation of undigested food take place. Then, from the sigmoid colon, digestive wastes move through the rectum and are egested at the anus.

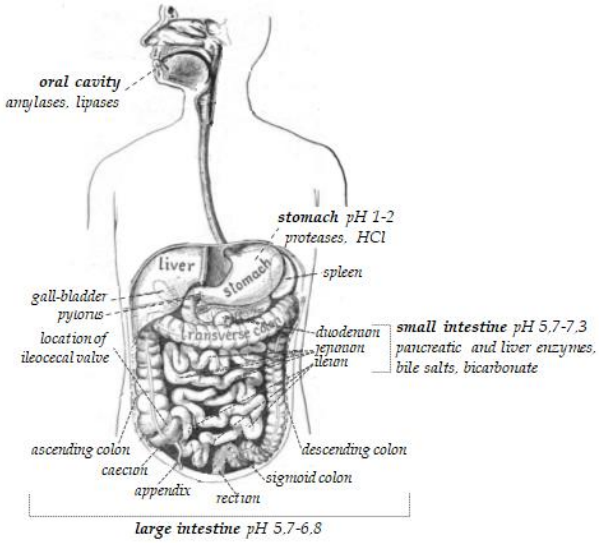


Figure 1: Human digestive system. *Functional regions with main digestive enzymes and pH (adapted from Sobotta's Atlas and Text-book of Human Anatomy, 1906)*

From the histological point of view the general structure of the GI tract (*Figure 2*), from outer layer towards the lumen, is composed of:

- **Serosa** - Loose connective tissue with elastic and collagen fibers, nerves and vessels, covered by a single layer of flat mesothelial cells. Where there is no mesothelial cover, the outermost layer is called **adventitia**.
- **Muscularis** - Composed of smooth muscle cells that are spirally oriented and divided into inner and outer sublayers.

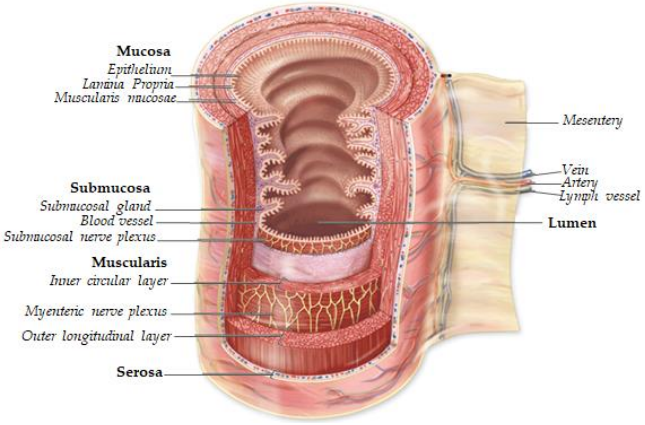


Figure 2: Major layers and organisation of the digestive tract *(adapted from Janqueira's Basic Histology, 2009).*

- **Submucosa** - Thick layer of denser connective tissue with numerous blood and lymphatic vessels.
- **Mucosa** (the main object of this manuscript) - The innermost layer that come in contact with the gastrointestinal content, further divided in different layers (from the outer to the inner layer): **muscularis mucosae**, a thin layer of smooth muscle separating the mucosa from the submucosa; **lamina propria** (LP) of loose connective tissue rich in blood and lymphatic vessels, lymphocytes and smooth muscle cells and the inner **epithelial layer**.

Despite this common histological organization, each intestinal segment is characterized by distinct histological characteristics, especially concerning the mucosa, that ensure the diverse functions performed along the GI tract.

1.1.1. Stomach

The stomach is a direct continuation of the oesophagus and can be divided in distinct regions: the most proximal part is the **cardia**, followed by the **fundus**, **corpus**, **antrum** and **pylorus** (*Figure 3*). The fundus and corpus constitute about 80% of the stomach and differs from the antrum and pylorus both functionally and histologically. At the gastro-esophageal junction, the mucosa abruptly changes from stratified squamous epithelium to simple cuboidal. The mucosa in the stomach is thick and lined with simple columnar epithelium. The surface epithelium invaginates into **gastric pits** into which the **fundus glands** open, producing and secreting the gastric juices (*Figure 3*).

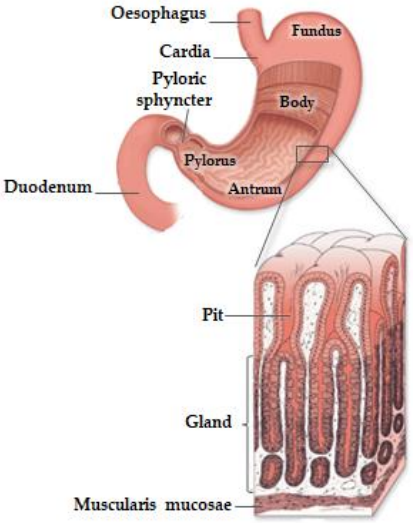


Figure 3: Regions and histological organisation of the stomach *(adapted from Janqueira's Basic Histology, 2009).*

1.1.2. Intestines

The small and large intestines, the main focus of this manuscript, form a continuous tube, lined internally by a single layer of columnar epithelium, stretches from the pylorus to the anus. Small and large intestine differ markedly in size, with the small intestine in human consisting of multiple coils amounting to 6-7m in length, whilst the colon is wider in diameter and much shorter (about 1,5m). However, both small and large intestines are devoted to the absorption of nutrients, electrolytes, vitamins and water.

Small Intestine

The small intestine is divided into three functional and histological distinct sections: duodenum, jejunum and ileum (*Figure 1*). Starting at the pylorus, the duodenum is the most proximal part of the small intestine and ends at the duodeno-jejunal junction. It receives partly digested food (chime) from the stomach and bile and pancreatic fluids from the pancreatico-biliary duct that neutralize the acidic content of the stomach. In the small intestine, secreted

pancreatic enzymes starts the digestion of lipids, carbohydrates and proteins to enable the absorption.

In order to increase the surface of absorption, the intestinal mucosa is highly folded, resulting in the formation of **crypts** and **villi** that extend into the intestinal lumen (*Figure 4*).

The absorptive and protective functions of the gut are dependent on an intact and functional intestinal epithelium, maintained by a constant cell renewal, starting at the bottom of the crypts where **stem cells** are located (*Figure 4*). From stem cells, several types of intestinal epithelial cells (**IECs**) originates, differentiating during the migration away from the replication zone and along the crypt-villus axis (*Figure 4*).

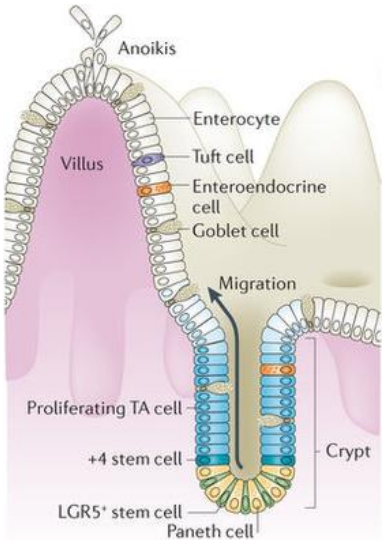


Figure 4: Organization of cell populations in the crypt and villi of small intestine *(Barker 2014).*

In close proximity to stem cells, the **Paneth cells** are situated in the crypts, producing antimicrobial peptides (AMP) and inflammatory cytokines (*e.g.* TNF-α, IL-17) (Porter et al. 2002), making them important for the formation of a chemical barrier and for the communication between the epithelial layer and the underlying immune system.

Located to the villi, columnar epithelial cells (**enterocytes**) guarantee the absorptive function. Interspersed between the enterocytes are **goblet cells**, specialized in the production of mucins, forming a **mucus layer** that cover the intestinal epithelium. The mucus layer consists of highly glycosylated proteins (mucins) forming a visco-elastic and gelly network that protects the intestinal mucosa. The thickness of the mucus layer increases proceeding from the duodenum towards the colon, although discontinuous in the small intestine.

Along the villi, a smaller population of **enteroendocrine** cells (including which D cells and L cells) produces a variety of hormones, among which peptide-YY (PYY) and glucagone-like peptide (GLP-1) from L cells and somatostatin from D cells, that regulates food intake, potentiates the glucose-induced secretion of insulin and control smooth muscle contractility, respectively (Batterham & Bloom 2003; Donnelly 2012; Huang et al. 1997).

Additionally, a recently characterized rare intestinal epithelial cell lineage is represented by the **Tuft cells**. These cells have been identified, along the villi, long ago (Isomäki 1973) but their function has remained unsolved until recently. Tuft cells have been recently described as

main source of endogenous intestinal opioids and the only epithelial cells expressing cyclooxygenase enzymes, suggesting a role in intestinal epithelium physiopathology (Gerbe et al. 2011). Additionally this cell lineage has been recently described as secreting IL-25 following helminth infection, thereby regulating type 2 immune responses in intestine (Gerbe & Jay 2016).

Interaction between IECs exist to maintain the intestinal barrier integrity and prevent leakage of large molecules between the cells into the lamina propria (LP). These interactions are assured by junctions controlling the paracellular flux between IECs (tight junctions) and regulating the communication and binding between cells (desmosomes, adhesion and gap junctions) (Garcia et al. 2017).

The intestinal mucosa is a selective barrier for nutrients, waters and ions. This permeability, fundamental for the absorption of nutrients, equally constitutes a vulnerability for infections. Thus, a specific, local innate and adaptive immune system resides in the intestine (detailed in section 1.2). Part of this immune protection is represented by clusters of subepithelial and lymphoid follicles constituting the **Peyer's patches** (*Figure 5*), dedicated to antigen sampling and initiation of the adaptive immune response. Peyer's patches are characterised by specialized epithelial cells called **M cells** (Microfold cells) compactly arranged with adjacent absorptive cells. In humans those cells do not have microvilli, but microfolds on the apical side of the membrane. M cells have been describe as capable of internalizing luminal antigens and presenting them to the lymphocytes thus playing a role in intestinal immune homeostasis (Neutra et al. 1996). Beside the Peyer's patches, **isolated lymphoid follicles** (ILFs) have been identified in the small intestine, functionally similar to the follicular unit composing the Peyer's patches (Hamada et al. 2002).

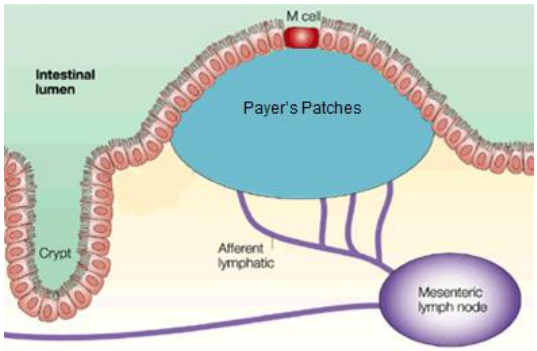


Figure 5: Localization of Peyer's patches and M cells in small intestine *(adapted from Mowat 2003)*

Large intestine

The large intestine is divided into: caecum, colon (ascending, transversal, descending and sigmoid), rectum and anal canal (*Figure 1*). Starting at the ileocecal valve, the large intestine is mostly devoted to the absorption of water, formation of the faecal mass from undigestible material, and production of mucus that lubricates the intestinal surface.

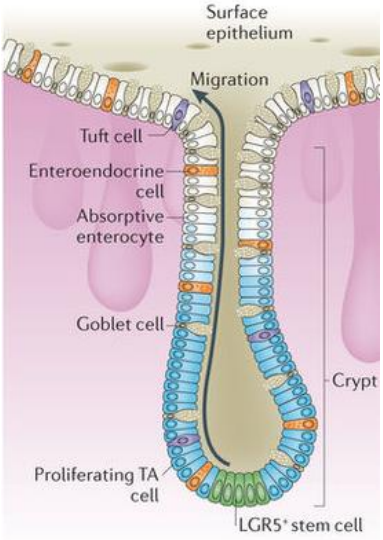


Figure 6: Organization of cell populations in the crypt of large intestine *(Barker 2014).*

The mucosa in the large intestine is smooth, without villi (*Figure 6*), contrary to the surface of the small intestine. However, it is penetrated throughout its area by tubular **intestinal glands** (crypts of Lieberkuhn). At the bottom of intestinal gland, **stem cells** are located, and originate several epithelial cell types, as in the small intestine (*Figure 6*). Intestinal glands are lined by absorptive and **goblet cells**, with a small number of **enteroendocrine cells**. The absorptive cells or **colonocytes** are columnar with short, irregular microvilli and constituting up to 80% of the intestinal epithelial layer. Along with their absorptive function, colonocytes secrete AMPs (β-defensins and cathelicidins) to overcome the absence of Paneth cells in the large intestine.

As in the small intestine, **goblet cells** are specialized in producing mucins that constitute the mucus layer covering the epithelia. Conversely to the discontinuous **mucus layer** characterizing the small intestine, the thickness of the mucus increases in the large intestine, becoming more compact in the colon. Here, the mucus layer is organized in two distinct areas: a dense and stratified inner layer covering the epithelium, and a loose outer layer in contact with intestinal lumen (Atuma et al. 2001). Additionally, moving from the intestinal epithelium to the intestinal lumen, the dense and stratified mucus layer becomes more loss due to a proteolytic dispersion of mucins (Atuma et al. 2001).

In the large intestine Peyer's patches are absent but equivalent M-cells containing macroscopic structures have been identified around the ileocecal valve (**caecal patches**) and through the colon and rectum (**colonic patches**) although their function still remains unclear (Owen et al. 1991). Consequently, the adaptive immune response in the large intestine, is mainly organized into **isolated lymphoid follicles** (ILFs). Additionally, the sampling of antigens from the

intestinal lumen is assured by a subset of dendritic cells (DCs) able to extend their dendrites between the tight junctions of the intestinal epithelial cells (Rescigno et al. 2001).

1.2. *The immune system in the GI tract*

The GI tract constitutes the largest surface of the human body exposed to the environment. Although the epithelia are covered by the mucus layer, the thin physical barrier is constantly challenged by potential harmful microbial components. Therefore, the intestinal mucosa is on the front line for **immunological defence**. A specific, local innate and adaptive immune system reside in the intestine to protect it against harmful microbial components. The organized structures of the **gut-associated lymphoid tissue** (GALT) and the draining lymph nodes are the principal locations for priming adaptive immune cell response in the intestine (*Figure 7*). Conversely, effector immune cells are diffusely distributed through the lamina propria and overlying epithelium.

The GALT comprises subepithelial lymphoid aggregates that lie in the mucosa and submucosa. Among these organized structures, **Peyer's patches** are characterized by an overlying-associated follicle epithelium consisting of numerous B cell lymphoid follicles, surrounded by small T cell areas. The Peyer's patches increase in density and size from the jejunum to the ileum and contain the **M cells**. M cells are specialized in the uptake and transport of particulate antigens from the lumen to an underlying **dendritic cells (DCs)-rich region** where they are presented to the adaptative immune cells.

In the colon, the adaptive immune response is mainly organized in **isolate lymphoid follicles** (ILFs) that, in contrast to Peyer's patches, consist of B cells with no clear T cell zone. As previously mentioned, the sampling of the intestinal lumen in colon is assured by DCs able to extend their dendrites between the tight junctions of the intestinal epithelial cells (Rescigno et al. 2001).

In order to avoid a systemic immune response to the constant bacterial threat in the intestine, immune responses in the GALT are primed in the **mesenteric lymph nodes** (MLNs), important site of T-cell activation for both the small and large intestine. DCs continually migrate from intestinal tissues to the MLNs where they present antigen and control the

development, migration, and functional differentiation of cells of the adaptive immune system.

Additionally, immunoregulatory signals are produced by **intestinal epithelial cells** (IECs) for tolerizing immune cells, limiting steady-state inflammation and directing appropriate adaptive and innate immune responses. Many of these responses are closely dependent on bacterial-derived components that drive a signalling from the IECs to the mucosal immune cells. Indeed, the production of the cytokines thymic stromal lymphopoietin (TSLP), transforming growth factor-beta (TGB-β) and IL-25 have been described as originated from bacterial-derived signalling. In particular, IECs-derived cytokines promote the development of DCs and macrophage with tolerogenic properties and drive the production of IL-10 (detailed in section 1.4).

Together with IECs, the mammalian intestinal epithelium contains numerous T cells that are located at the basement membrane between enterocytes at a frequency of up to 10-15 **intraepithelial lymphocytes** (IELs) per 100 epithelial cells, forming one of the main branches of the immune system (Ferguson 1977). IELs serve as a gateway guards at the intestinal epithelial surface constantly monitoring the epithelial layer for infection or disruption. Two major groups of lymphocytes that reside within the intestinal epithelial layer can be distinguished based on expression of either a αβ T-cell receptor (TCR) or a γδ TCR (Sheridan 2011). In humans, TCRγδ$^+$ IELs represent about 10% of the small intestinal IEL, but drastically increase under certain allergic or inflammatory conditions (Spencer et al. 1991).

A growing family of immune cells are the **innate lymphoid cells** (ILCs), found at the barrier surfaces, such as the intestine, where they function as regulators of tissue homeostasis, inflammation and early innate immune response to infection (Tait Wojno & Artis 2012). ILCs are regulated, in part by epithelial-cell derived immune-regulatory signals such as IL-25 and IL-33 both stimulating and suppressing different groups of ILCs (Koyasu & Moro 2012). ILCs are extremely heterogeneous and the various subset are differently distributed in the epithelium of small and large intestine, probably influenced by the distinct digestive functions. ILCs are characterized by their developmental requirements and differential cytokine expression into group 1, group 2 and group 3.

Group 1 ILCs includes natural killers (NKs) and innate lymphoid cell subset 1 (ILC1), characterized by the production of the T_H1 cell-associated cytokine interferon-γ (IFN γ) and tumor necrosis factor (TNF) in response to IL-12 and/or IL-15 (Bernink et al. 2013).

Group 2 ILCs (ILC2), produces the T_H-2 associated cytokine IL-5 and IL-13, contributing to an early innate response to intestinal helminth infection and invoke a protective epithelial response including the stimulation of goblet cells in increasing mucus secretion (Neill et al. 2010). Additionally, considering the role of ILC2 in hyper-responsiveness and tissue repair in lung following allergy or influenza infection (Monticelli et al. 2011; Halim et al. 2012), a similar role was proposed in the intestine perhaps during food allergy. The proliferation and activation of ILC2 is supported by predominantly epithelial cell-derived cytokines IL-25, IL-33 and thymic stromal lymphopoietin (TSLP) (Camelo et al. 2017).

Group 3 ILCs produces Th17 and Th22-associated cytokines, including IL-17A and IL-22 in response to stimulation with IL-23. This group includes ILC3s and lymphoid tissue inducers (LTi), involved in secondary lymphoid tissue organogenesis. ILC3-derived IL-22 has been reported as important in protecting the intestinal epithelium from injuries as well as pathogen infections (Zenewicz et al. 2008; Sonnenberg et al. 2012). Conversely, ILC3- derived IL-17 is thought to have a pro-inflammatory effect in the intestine, implicated in both murine colitis and human inflammatory bowel diseases (IBD) (Buonocore et al. 2010; Geremia et al. 2011).

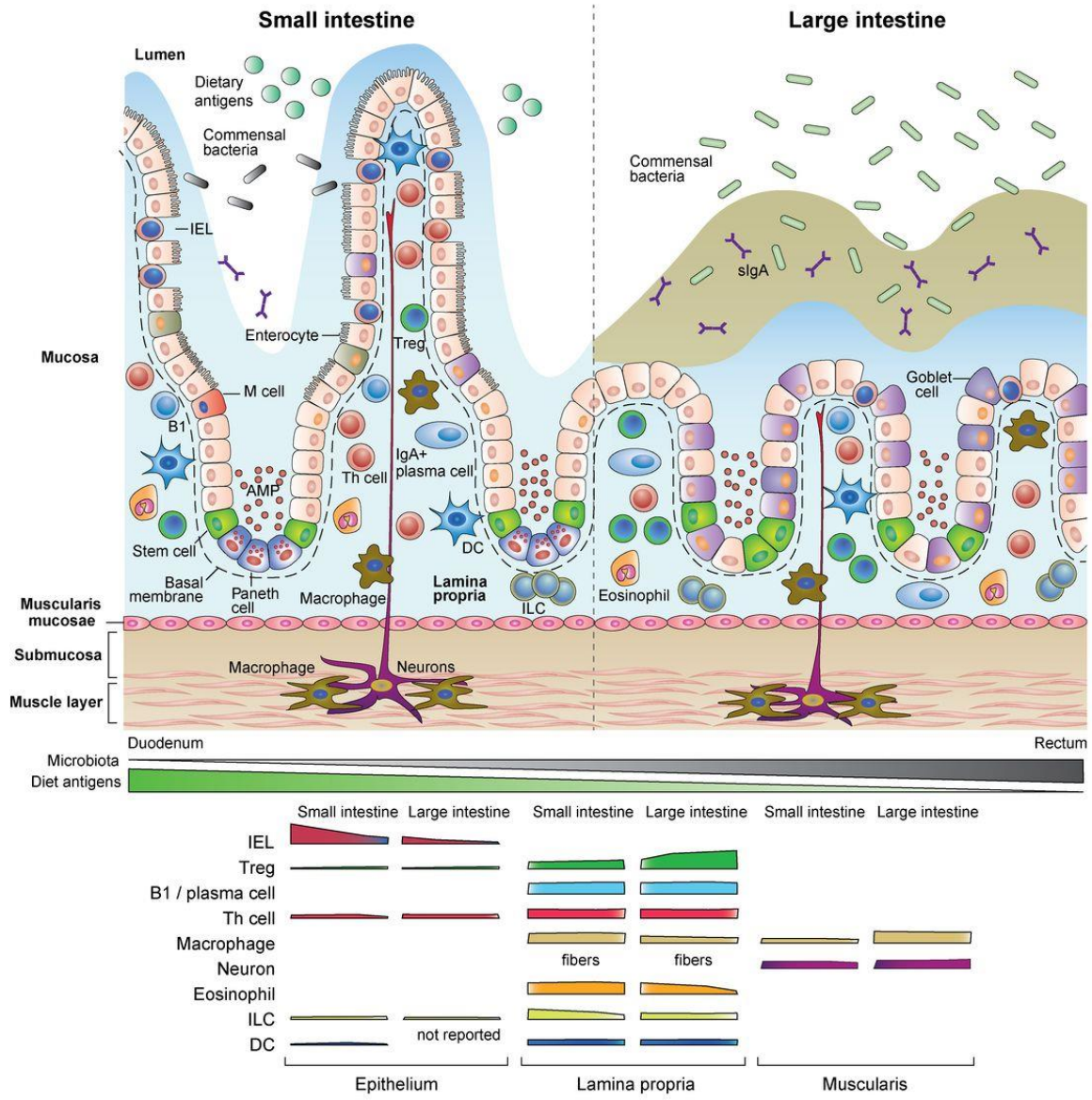


Figure 7: Intestinal microenvironment and niches. *Main characteristics of the intestinal layers, proximal-distal regions and immune cell population exposed to chronic stimulation by dietary and microbial antigens. Although stimulation by dietary antigens or metabolites decreases from proximal to distal intestine, microbial stimulation follows the opposite direction. An approximate illustration of the changes in abundance of each cell type per intestinal region is shown on the bottom (Faria 2017).*

1.3. *The Human Gut Microbiota*

1.3.1. Composition of the gut microbiota

During lifetime the GI tract enters in contact with food as well as with an abundance of microorganisms, some of them colonizing the intestinal epithelium which therefore constitutes a large ecological niche. The human GI ecological niche harbours 10^6–10^{14} microorganisms belonging to eukaryotes, prokaryotes and virus, collectively referred as gut microbiota. The physico-chemical characteristics of each anatomical segment along the GI tract affect the bacterial community. Microbial density and diversity increase gradually from the stomach to the colon, influenced by different factors among which the pH, redox potential and transition time (Donaldson et al. 2016) (*Figure 8*).

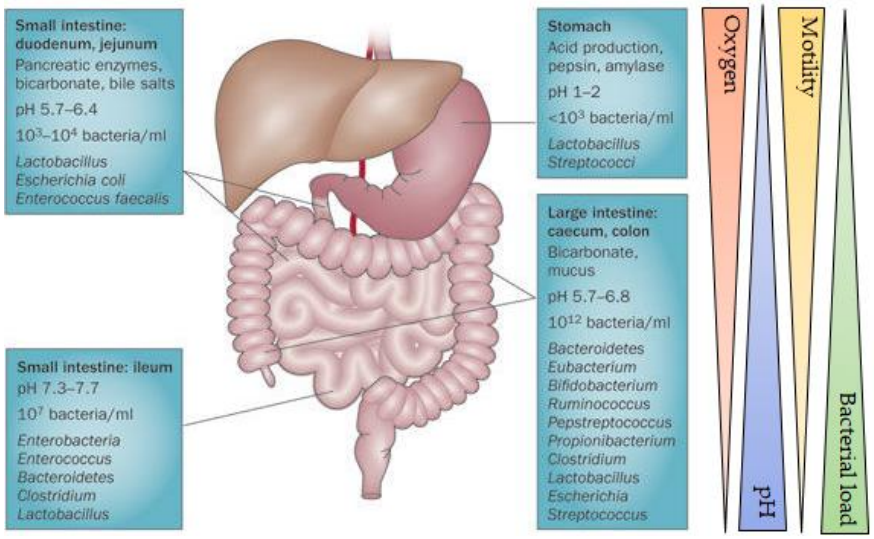


Figure 8: Characteristics of the normal gastrointestinal tract. *The various organs of the gastrointestinal tract differ according to digestive secretions and pH. Different species and quantities of bacteria are found at different points along the digestive tract according to these major variations in the environmental niche (pH, Oxygen, Motility) (Aron-Wisnewsky 2012)*

The small intestine is typically more acidic and has higher levels of oxygen and antimicrobials than the colon. Therefore, the small intestine is dominated by fast-growing facultative anaerobes that tolerate the combined effects of bile acids and antimicrobials. Additionally, the shorter transit time in the small intestine compared to colon further selects the resident bacteria for their capacity to adhere to mucus or epithelia (Donaldson et al. 2016). In ileostomy patients it was evidenced a lower bacterial diversity in the small intestine than the colon, although the

resident facultative anaerobes are more specialized in the import of simple sugars (Zoetendal et al. 2012). In contrast, the colonic conditions support a more dense and diverse microbial community of bacteria, mainly anaerobes, specialized in the degradation of complex undigested carbohydrates (Donaldson et al. 2016).

Additionally, within the colon, bacteria are also organized along the transverse axis, from the middle of the lumen towards the mucosa. Examination of the colon by fluorescence *in situ* hybridization (FISH) has shown that the inner mucus layer appears essentially sterile next to the densely populated outer layer (Johansson et al. 2008) (*Figure 9*). Coherently, it has been evidenced that the outer mucus of the large intestine forms a unique microbial niche in which bacterial species show differential proliferation and resource utilisation compared with the same species in the intestinal lumen (Li et al. 2015). Bifidobacteria (including *Bifidobacterium breve* and *Bifidobacterium longum*), Bacteroides (including *Bacteroides thetaiotaomicron*), *Ruminococcus* genera and *Akkermansia muciniphila,* possess the ability to degrade and adhere to the complex sugar structure of mucins (Derrien et al. 2008; Marcobal et al. 2013). Beside these, the mucosal niche allows the formation of obligate mutualistic metabolisms within bacteria able to use mucins as carbon source and those only able to adhere but not degrade the polysaccharides, such as *Escherichia coli*.

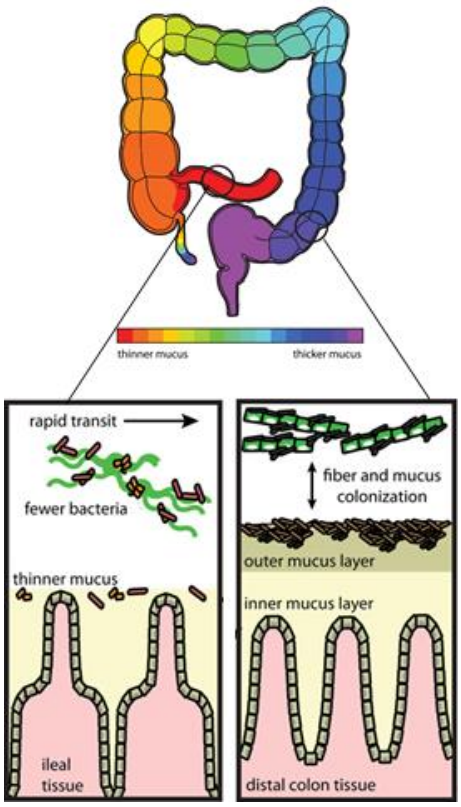


Figure 9: Spatial distribution of microbiota in the mucus layer of small and large intestine in relation to mucus thickness (*colours from red to purple indicate increasing mucus thickness) (adapted from Koropatkin 2012)*

Despite the variations of the bacterial composition and density among different body sites, it has been revealed that human-associated microbial communities contain sufficient strain-level variation to discriminate between individuals in a fixed population over time (Franzosa et al. 2015). Although, a limited number of dominant and prevalent bacterial species was evidenced as shared among individuals (Tap et al. 2009). Consequently, the definition of a **microbiota phylogenetic core** was proposed to represent this shared microbial community. Identified

from faecal samples of healthy humans, the core microbiota was described as dominated in bacteria belonging to Bacteroidetes and Firmicutes, with a reduced abundance in Actinobacteria, Proteobacteria and Verrucomicrobia (Tap et al. 2009; Turnbaugh et al. 2009). In line with the identification of an inter-individual microbial core, the composition of gut microbiota in healthy human is referred to be quite stable during adult life (at least up to 65 years old) (Lozupone et al. 2012). To shed light on the individual gut microbiota, human faecal samples from three countries (Denmark, Spain and the United States) have been clustered based on their taxonomic composition and resulted in the proposal of three "densely populated areas in a multidimensional space of community composition", termed **enterotypes** (Arumugam et al. 2011). The classification of the gut microbiota in three different enterotypes, independent of age, gender, cultural background and geography, is based on the prevalent phyla in the studied population (Arumugam et al. 2011). A single bacterial genus dominates two of three described enterotypes (*Bacteroides* or *Prevotella*) while a fairly limited set of *Ruminococcaceae* is marker for the third enterotype (Arumugam et al. 2011) (*Figure 10*).

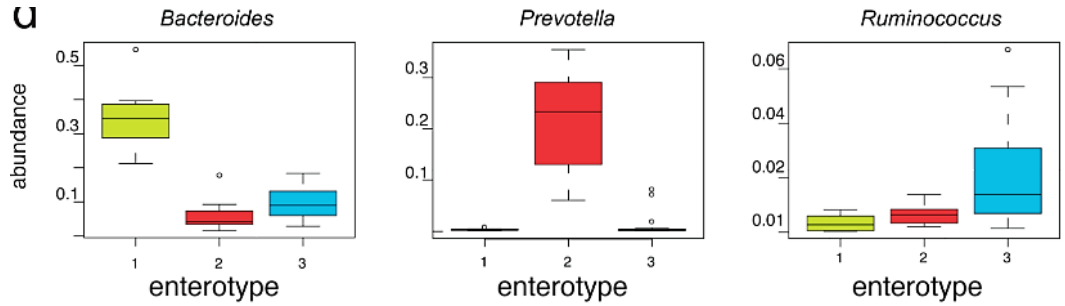


Figure 10: Phylogenetic differences between enterotypes. *Abundance of the main contributors of each enterotypes (Arumugam 2011).*

Recently, an alternative approach for analysing the microbial community in different cohorts, let to the identification of four groups, instead of the originally proposed three. Two of these overlapped the *Bacteroides* enterotype and the *Prevotella* enterotype, while the other two seems a more complex mixture (Costea et al. 2017) (*Figure 11*). Albeit highly debated, the definition of enterotype clusters represent a useful tool to reduce the complexity of the intestinal ecology and easily analyse the effects of microbial population on host's physiology.

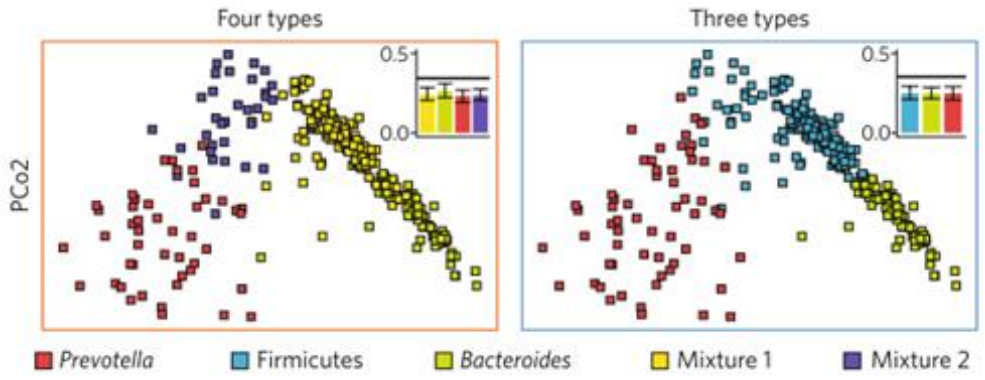


Figure 11: Stratification of the microbial composition landscape of the human gut microbiome. *Projection onto a set of Danish samples of the three most frequent enterotype classification schemes defining four groups based on genera abundance (adapted to Costea 2017)*

Even if the gut microbiota composition is defined as quite stabile during adult age, enterotypes were originally defined as not fixed. In fact, the number and composition of the bacterial community is altered by the intestinal and host's environment during lifetime (*e.g.* age, diet, antibiotic consumption and geography) (Cotillard et al. 2013; Biagi et al. 2010). Indeed, 5% of the inter-individual variability is defined by dietary habits (Tap et al. 2009) and consequently, also the enterotypes are influenced by diet (Wu et al. 2011).

Difference in the taxonomic composition of the enterotypes suggests distinct functional and ecological properties with consequent alternative utilization of components in ingested food. Accordingly, *Bacteroides* entetotype is favored by a diet rich in animal proteins and fats, typical in western societies, in line with a large proportion of *Bacteroides*-specific carbohydrate-active enzymes (CAZymes) (Cantarel et al. 2009) specialized for animal carbohydrates utilization (Costea et al. 2017). Conversely, the *Prevotella* enterotype would be most prevalent in people with high fiber diets, rich in fruits and vegetables, in line with the specialization of *Prevotella* hydrolases for the degradation of plant fibres (Purushe et al. 2010). Coherently, a study on children from Burkina Faso and Italy highlighted extremely marked differences attributed to dietary habits and the geographic localisation, with significant abundance in *Prevotella* and *Xylanibacter* and concomitantly depletion in *Firmicutes* and *Enterobacteriaceae* (*Shigella* and *Escherichia*) in African child compared to the Europeans, due to the fiber-rich diet (De Filippo et al. 2010). Accordingly, the abundance in *Bacteroides* spp., *Bifidobacterium* spp., *Escherichia coli* and *Enterobacteriaceae* spp. is significantly lower in faecal samples from vegetarians compared to omnivores (Zimmer et al. 2012). Moreover, a diet rich in *Brassica* vegetables evidenced, in

human, a reduced abundance of sulphate-reducing bacteria and members of *Rikenellaceae*, *Ruminococcaceae*, *Mogibacteriaceae*, *Clostridium* and unclassified *Clostridiales* (Kellingray et al. 2017).

Although host genetics and geographic affiliations have an impact on microbial composition, lifestyle has a major effect on the bacterial community. Indeed, beyond the diet, therapeutic treatments such as chemotherapeutic agents and antibiotics (also referred as xenobiotics) are described as important modulators of the microbiota composition. The administration of amoxicillin/clavulanic acid during 10 days was reported to cause a post-antibiotic diarrhoea and a decrease in *Clostridium leptum*, *Clostridium coccoides* and Bifidobacteria (Young & Schmidt 2004). Moreover, upon discontinuation, the original taxonomic composition was only partly recovered with taxa still absent after 4 weeks (Dethlefsen & Relman 2011), suggesting a long-term impact of antibiotics on gut ecology.

1.3.2. Development and age-related variation of gut microbiota

Despite the variations due to environmental factors, the microbiota of healthy adults, could be still defined as quite stable. However, during lifetime from newborn to elderly, the microbial density and composition evolve dramatically (*Figure 13*).

The colonization of the intestinal tract starts very early in the animal life. Recent studies suggested that the bacterial colonization of the newborn already occurs in uterus; indeed the microbiota composition in infant meconium shares features with the bacterial communities identified in placenta and in amniotic fluid, suggesting a prenatally gut colonization (Collado et al. 2016). At birth, the neonatal colonization by microbiota is further influenced by a variety of environmental factors, including the mode of delivery, maternal microbiome and the hygiene of the neonatal environment. Then, the link between the mother and the offspring perpetuates after birth with microbes present in breast milk. Coherently, at the age of 3–4 days, the infant gut microbiota composition is similar to the bacterial community present in colostrum (Collado et al. 2016). Neonatal are rapidly colonized by facultative anaerobic-aerobic bacteria (Staphylococci, Enterococci, Enterobacteria) through the mother and the environment. However, these bacteria, reaching high densities in few days, consume the

oxygen in the intestinal tract, with a consequent implantation of strict anaerobic genera such as *Bifidobacterium*, *Bacteroides* and *Clostridium*.

The newly implanted microbial community diversifies during the first months of life and then stabilizes around the age of 3 years, when its composition resembles the microbiota of the adult (Yatsunenko et al. 2012). However, the adult microbiota is likely to maintain the imprinting of the infant colonization, primarily influenced by the mother, method of delivery and breast feeding. Indeed, low level of Bifidobacteria have been observed in babies whose mother gained significant weight during pregnancy as well as in asthmatic patients whose mother have significantly reduced Bifidobacteria species in her breast milk (Grönlund et al. 2007).

The relative stable adult microbiota undergoes a major change in composition as well as in diversity with ageing, characterized by lower Firmicutes/Bacteroidetes ratio in adults over 65 compared to younger adults (Biagi et al. 2010) (*Figure 12*). However, conflicting results emerged when comparing different cohorts, probably influenced by the reduced bacterial diversity and the different geographical localization (Mueller et al. 2006; Biagi et al. 2010). For example, a small percentage of Firmicutes was observed in Japanese, Finnish and Italian population whilst in Germans the trend was reversed (Mueller et al. 2006). Conversely in an Italian study, young and elderly adults showed a very comparable overall structure of the gut microbiota, whilst centenarians emerged as a separate population, with a significant lower density and a different composition than the adult-like pattern (Biagi et al. 2010). Specific changes in relative proportion of Firmicutes subgroups are observed in extreme old people, with a decrease in the contributing *Clostridium* cluster XIVa, an increase in Bacilli, and a rearrangement of the *Clostridium* cluster IV composition. Moreover, the gut microbiota of centenarians was also described as enriched in *Proteobacteria* that, under some circumstances (*e.g.* inflammation), could induce pathologies (Biagi et al. 2010).

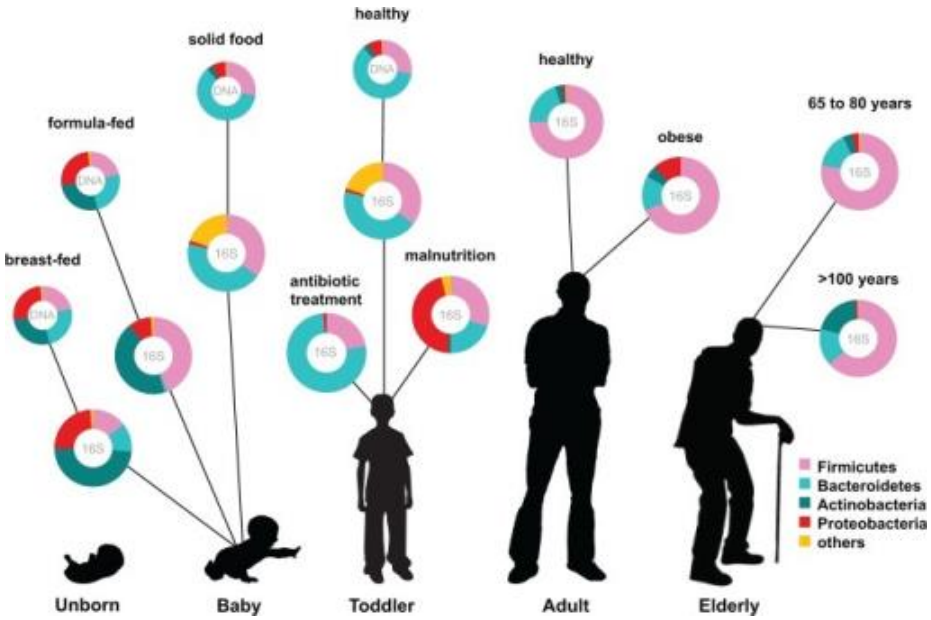


Figure 12: Human microbiota composition through life stages. *The graph provides a global overview of the relative abundance of key phyla of the human microbiota composition in different stages of life (Ottmann 2012).*

Despite the overall variations in the predominant Bacteroidetes and Firmicutes phyla during lifetime, important oscillations in bacterial composition have been observed also at level of microbial genera. In particular, Bifidobacteria genus, belonging to Actinobacteria phylum, dominates the intestine of healthy breast-fed infants and decrease along with age (*Figure 13*). *Bifidobactreium longum*, *Bifidobacterium breve* and *Bifidobacterium bifidium* are generally dominant in infants with differences between brest- and formula-fed infants (*e.g. B. longum subsp. infantis* and B. *longum subsp. longum* associated to brest-and formula-fed respectively) (Klaassens et al. 2009). Conversely, *Bifidobacterium catenulatum*, *Bifidobacterium adolescentis* as well as *Bifidobacterium longum* are prevalent in adult (Gavini et al. 2001). However, in adulthood the density of Bifidobacteria are stable but lower compared to the infant (*Figure 13*).

The decline in bacterial diversity observed at phylum level along ageing, was also reported in Bifidobacteria, associated with a reduction in the adhesion capacity to intestinal mucosa (*Figure 13*). In centenarians, the characterization of Bifidobacteria diversity and composition seem controversial. Indeed, *Bifidobacterium longum* has been reported as the most abundant in Italian centenarians followed by *B. adolescentis* and *B. bifidum* (Biagi et al. 2010), while *B. dentium* was dominant in Chinese centenarians (F. Wang et al. 2015).

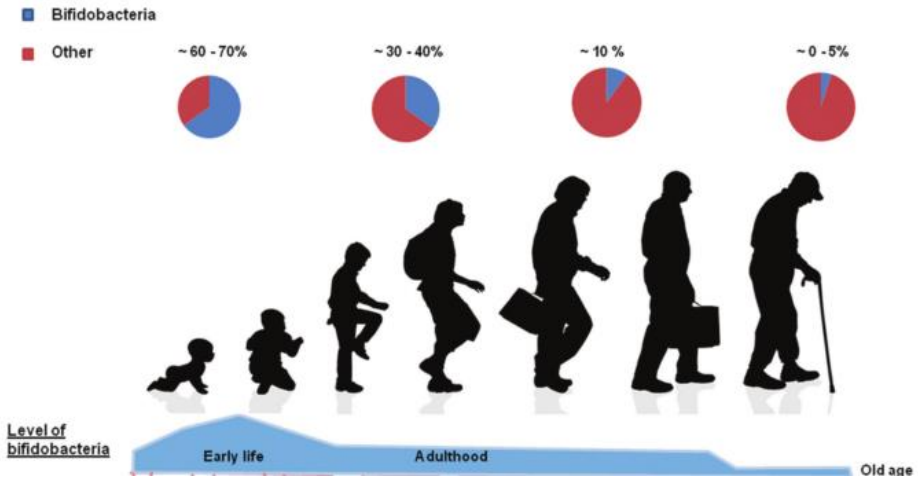


Figure 13: Gut Bifidobacteria composition through life stages. *The graph provides a global overview of the relative abundance of Bifidobacteria in human microbiota in different stages of life (Arboleya 2016).*

In conclusion, microbiota adapts to the environmental changes during lifetime, responding mainly to lifestyle and host's genetics. However, the divergence in reported microbial compositions of different cohorts, reflects the difficulties to access the gut ecological niche and the technical variables occurring from the collection to the analysis (among which the collection method, time and conservation, and the analysis of the collected samples) likely to affect the final results.

1.3.3. The study of gut microbiota

The extended knowledge on gut microbiota, took the advantage from culture-independent molecular approaches. Indeed, the dominant human fecal microbiota is considered not completely exploitable using classical microbiological techniques. This promoted the development of comparative sequencing of 16S rDNA amplicons and more recently sequencing of the combined genomes of all dominant microbes within a given ecosystem, termed the metagenome.

The first extensive gene catalogue of human metagenome encompassed 3.3 million non-redundant genes from a cohort of 124 European individuals (Qin et al. 2010), then expanded to 1267 subjects, including Americans of the Human Microbiome Projects as well as Chinese individuals. The metagenome catalogue gave access to the discovery of new microbial genes, involved in host's physiology. Interestingly, some of the identified genes not concerned the

core microbiome, but rather consisted of rare genes that are found in a limited fraction of the population (Qin et al. 2010), confirming that the molecular functions are not necessarily provided by abundant microbial species.

Despite the huge effort in the field of metagenomics to deeply decipher the gut microbiota, it recently emerged that cultivation and molecular methods are complementary. Indeed, strains that can be cultured *in vitro* are currently indispensable for the description of novel diversity and eventually the improvement of taxonomic and sequence databases. Additionally, the cultivation approach allows the use of minimal consortia of microbes, helpful for detailing and standardizing studies of gut microbial communities and microbe-host interactions. Besides, molecular techniques are helpful as they provide insights into strain-level diversity and the functional potential of organisms. Furthermore, genomic and metagenomic data allow inferring growth conditions for uncultured bacteria and enable detailed genetic studies.

Consequently, the term 'culturomics' was created, referring to a large-scale cultivation of gut bacteria on agar plates followed by high-throughput mass spectrometry-based identification and genome sequencing (Lagier et al. 2012). In line with this, a recent described microbiological approach coupled to genomics, isolated 137 bacterial species from the human gut, including 67 novel taxa (Browne et al. 2016), thus confirming the huge potential of a coupled approach for the characterization of the resident microbiota.

Consequently, the extended knowledge on the complex composition of gut microbiota and the variety of microbial (potential) metabolisms occurring in the intestinal tract, prompted a raised considerable interest on host-microbiota dialog to identify the influence of intestinal bacteria on host's physiology, by taking the advantage of emerging omics tools (metagenomics, proteomics, genomics).

1.4. *The Role of Gut Microbiota in the Host-Bacteria Cross-Talk*

The GI tract is a critical interface between human cells, the external environment and the commensal bacteria, where, both the host and microbes co-evolved for millions of years, establishing a fine regulated cross-talk. However, the impact of colonic bacteria is not limited locally on the colon, as the microbiota is capable of affecting gene expression and biological function even in more distant organs, including liver, pancreas and the brain. Indeed, microbiota has been described as implicated in brain function (gut-brain axis), liver function (gut-liver axis), mucosal and systemic immune function, diabetes (type 1 and type 2) (gut-pancreas axis), obesity, and cardiovascular diseases.

1.4.1. The Barrier Homeostasis in Intestine

Despite its constant exposure to a high dense microbial community, the intestinal immune system is in a "homeostatic" state with the local microbiota, in healthy adults. This means that an equilibrium set point is maintained through positive and negative host responses evolving simultaneously with external conditions and the microbiota. This intestinal immune homeostasis is principally guaranteed by biological barriers.

The integrity of the **epithelium** constitute a first physical barrier, guaranteed by the constant renewal of epithelial cells (refer to GI section). Microbiota is also concerned in the maintenance of this physical barrier by modulating the genes involved in tight junctions formation and in epithelial repairs. Indeed, a commensal strain of *Lactobacillus reuteri,* isolated from pig intestinal lumen, was reported as an inducer of cellular proliferation in intestinal organoids and *in vivo* (Hou et al. 2015), thus likely to protects the intestinal barrier. Moreover, it was evidenced that muramyl-dipeptide (MDP), a peptidoglycan motif common to all bacteria, triggers stem cell survival in murine organoids, through the stimulation of the cytosolic innate immune sensor Nod2 (Nigro et al. 2014). However, the thin epithelial layer is a vulnerable barrier, so the mucus layer constitutes an additional protection barrier.

The **intestinal mucus**, stratified between the deep area, made of a virtually sterile compact mucus, and a superficial area, colonized by bacteria (refer to section 1.3) (Zoetendal et al. 2002), represents a first physical barrier to protect intestinal epithelium. Indeed, mice deficient for *Muc2*, a gene encoding the main intestinal mucin, present spontaneous colonic inflammation

due to the lack of the bacterial-free area in the mucus stratification (Van der Sluis et al. 2006). In turn, the production of mucus by goblet cells (refer to section 1.1), is stimulated by gut microbiota. Indeed, in the absence of bacteria, the number and size of goblet cells appears reduced and the thickness of the mucus layer decreases, compared to animal with conventional microbiota (Kandori et al. 1996) (*Figure 14*). In addition to the mucus density serving as a physical barrier for microorganisms, antimicrobial peptides (AMPs) and immunoglobulins, secreted by the epithelium, accumulate in higher concentration in the mucus, further controlling the microbial population (Donaldson et al. 2016) (*Figure 14*).

The **antimicrobial peptides** (AMPs), secreted by intestinal epithelial cells (IECs) and Paneth cells (refer to section 1.1), are a different class of bactericidal molecules including defensins, cathelicidins and type C-lectins, with a broad antibiotic activity (Hooper & Macpherson 2010). Some of them are constitutively expressed, such as the α-defensins, while others depends on microbial signals via the stimulation of immune receptors termed Pattern Recognition Receptors (PRR) (further detailed in section 1.4.2) (Hooper & Macpherson 2010). These bactericidal molecules, accumulated in the mucus layer, form a gradient for an enhanced protection of the epithelial cells without influencing the composition of the luminal microbiota (Kobayashi et al. 2005).

A third mechanism reported to sustain the intestinal physical barrier is the secretion of the antibody **IgA** (Immunoglobulin A) (*Figure 14*). Sampling of the intestinal microbiota by dendritic cells (DCs) leads to the differentiation of B cells into IgA-producing plasma cells. IgA and IgM are the only antibody isotypes able to reach the intestinal lumen and therefore prevents association of bacteria to the epithelium (Shulzhenko et al. 2011).

Despite this physical barrier, it was reported that a small amount of intestinal microorganism escape this control and penetrate through the epithelial monolayer (Shanahan 2002). Therefore, the immune system associated with the gut guarantees a second line of defense (refer to section 1.2). Intestinal macrophages in the lamina propria eliminate intruding bacteria by phagocytosis and lysosomal degradation. The maintenance of the balance between sufficient immune responses to invasion and a strictly local response to the frequent events of microbe-host crosstalk, is guaranteed by the organized structures of the gut-associated lymphoid tissue (GALT –refer to section 1.2). Complementing these local barriers, a third and crucial mechanisms take place for the maintenance of the homeostasis. The minimization of

exposure of resident bacteria to the systemic immune system is guaranteed by the mesenteric lymph nodes which allow the priming of the mucosal tissue while preventing the immune response to become systemic (MacPherson & Uhr 2004).

The co-evolution of eukaryotic cells in close proximity with microbial community, shapes a variety of sensing mechanism to discriminate commensal bacteria from potential harmful microbes.

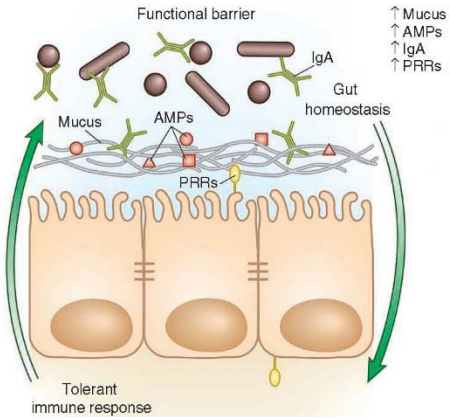


Figure 14: The regulation of Intestinal Barrier to ensure a tolerant response to microbiota *(adapted from Brown 2013)*

1.4.2. Mechanisms for Sensing Microbial Signals by IECs

The maintenance of barrier and immunoregulatory functions, is guaranteed by the ability of intestinal epithelial cells (IECs) to act as a frontline sensor of microbes and to integrate commensal bacteria-derived signals into antimicrobial and immunoregulatory responses.

IECs express **pattern-recognition receptors** (PRR) that enable them to act as a dynamic sensor of the microbial environment and as active participants of the mucosal immune response. The PRRs sense evolutionarily conserved microbe-associated molecular patterns (MAMPs) of microorganisms (*e.g.* flagellin, peptidoglycan, lipoteicholic acid) that trigger sequential activation of intracellular signalling pathways, leading to the induction of cytokines and chemokines to modulate the early host resistance to infection. Members of PRR include Toll-like receptors (TLRs) and Nucleotide binding Oligomerization Domain (NOD)-like receptors (*Figure 15*).

Toll-like receptors (TLRs) (*Figure 15*) are PRRs expressed by various cells in the gastrointestinal tract, including intestinal epithelial cells (IEC) and resident immune cells in the lamina propria. TLR signalling is involved in either maintaining intestinal homeostasis or the induction of an inflammatory response. TLRs recognize a wide range of microbial fragments and therefore sense antigens derived from both the microbiota and from invading pathogens. TLR2, dimerizing with TLR1 or TLR6, recognizes bacterial cell wall lipoproteins. Lipopolysaccharide (LPS) produced by Gram-negative bacteria is recognized by TLR4, whereas flagellin is recognized by TLR5. In addition, bacterial DNA is recognized by TLR9. Coherently with their role, TLR2, 4, and 5 are generally expressed at the cell membrane, whereas TLR9 is expressed intracellularly. However, in IEC, TLR9 has been reported to be also expressed at the cell membrane (Lee et al. 2006). Under homeostatic conditions, IEC show low expression of TLR2 and TLR4 and therefore they are unresponsive to their TLR stimuli (Otte et al. 2004). Additionally, a bacterial polysaccharide (PSA) produced by *B. fragilis* activates TLR2 directly on Foxp3+ regulatory T cells to induce mucosal tolerance (Round et al. 2011). However, under inflammatory conditions, epithelial TLR expression is increased, which contributes to both inflammation as well as immune tolerance (Otte et al. 2004). The TLRs signalling, dependent on two adaptor molecules MyD88 (all TLRs except TLR3) and TRIF (TLR3 and TLR4), induces the production of antimicrobial peptides (AMP) belonging to the C-type lectin family (*e.g.* Reg3β and Reg3Ɣ), in response to bacterial signals (Kawasaki &

Kawai 2014). Thus, the fine regulated cross-talk between the host and the microbiota, is also mediated by the TLR signalling. Indeed, MyD88$^{-/-}$ mice are associated with both a shift in bacterial diversity and a greater proportion of segmented filamentous bacteria (SFB) in small intestine (Larsson et al. 2012) that could have potential pathogenic roles. It has recently been reported that the expression and secretion of soluble protein of the lectin family (galectin-9) is supported by the TLR9 activation, inducing tolerogenic dendritic cells (DC) along with the development of the Treg cells (de Kivit et al. 2013). Furthermore, a recent study demonstrated that *Clostridium butyricum* increased iTreg generation via a TLR2-dependent induction of TGFβ1 by DCs (Kashiwagi et al. 2015). This TLR2-dependent generation of tolerogenic DC favours regulatory T-cells induction which is similar to what was reported for *Bacteroides fragilis* (Round et al. 2011) suggesting a more general mechanism used by several commensal bacteria.

In addition to its regulatory role on the intestinal immunity, TLR activation on IEC is also known to modulate the expression of tight junction proteins. In particular, epithelial TLR2 activation has been described to protect against barrier disruption by enhancing tight junction expression in IECs in a protein kinase C-dependent manner (Cario et al. 2004). In contrast, the activation of TLR4 increases intestinal permeability and results in enhances bacterial translocation (Li et al. 2013) likely through NFκB signalling activation, that in turn is suppressed by apically TLR9 activation with a consequent prevention of TLR4-induced gut leakiness (Lee et al. 2006).

Another class of PRR activated by commensal-derived metabolites are the cytosolic **NOD-like receptors** (NLR) (*Figure 15*). NLR family members are cytosolic sensors crucial in the intestinal innate immune system by controlling both the commensal microbiota as well as enteropathogenic bacterial infections. In particular NLR are described to mediate the activation of Caspase-1 and the subsequent processing and secretion of the pro-inflammatory cytokines IL-1β and IL-18. The pattern recognition receptors NOD1 and NOD2 are among the best-studied NLRs, and their ligands have been extensively described. Both NOD1 and NOD2 sense cytosolic bacterial peptidoglycan fragments with high specificity: NOD1 is activated by D-glutamyl-meso-diaminopimelic acid (DAP) containing peptidoglycan fragments, which are mainly found in Gram-negative bacteria, whereas NOD2 was shown to bind and responds

to muramyl dipeptide (MDP), found in all bacteria (Girardin, Ivo G. Boneca, et al. 2003; Girardin, Ivo G Boneca, et al. 2003).

Consistent with a role of NOD1 and NOD2 in host responses against bacterial infection, *Nod2*$^{-/-}$ mice showed an altered microbiota and an enhanced susceptibility to several pathogens (Couturier-Maillard et al. 2013). Coherently, patients with Crohn's disease, from whom NOD2 gene polymorphisms are observed, have altered composition of ileal microbiota regardless of their inflammatory status (Rehman et al. 2011).

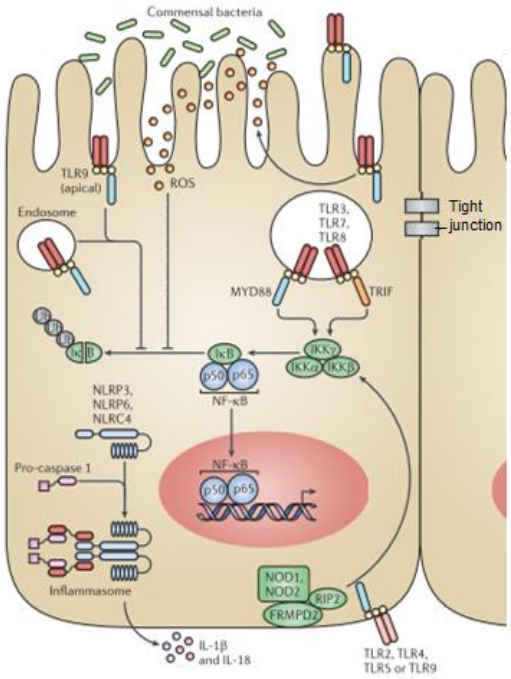


Figure 15: IECs sense microbial signals. PPRs including TLRs and NLRs recognize MAMPs inducing effects on barrier integrity *(adapted from Peterson 2014).*

In conclusion, the maintenance of intestinal homeostasis could not be considered exclusively dependent on the host sensing capacity but also on a concerted action of bacterial and hosts responses.

1.4.3. Intestinal Microbiota and the Immune System beyond the Gut

While the link between microbiota and the local mucosal immune response has been largely described, more limited is the knowledge on the impact of commensal bacteria on peripheral responses. The development of antibiotics and the improvement of hygiene have led to a significant reduction in infections but also to an increased susceptibility to autoimmune and allergic diseases (Russell et al. 2012; Bach 2002). For the human population, antibiotics are seen as a major modifiers of the beneficial human-microbiota crosstalk, together with alterations caused by other exogenous factor, such as urbanization, global travel and dietary changes (Dethlefsen et al. 2007). In experimental models, antibiotic administration modifies microbiota structure and is linked to an increase susceptibility, for example, to allergic airway inflammation and food allergies (Bashir et al. 2004; Noverr et al. 2005). The long-term consequences of microbial perturbation through the intensive use of antibiotics are difficult to discern, although chronic conditions, such as asthma, have been associated with childhood antibiotic use and altered microbiota.

More direct evidences have been provided for the role of microbiota on peripheral responses and in particular the innate immune cell development. Germ-free (GF) mice display reduced proportions and differentiation potential of myeloid cell progenitor populations of both yolk sac and bone marrow origin. The defect in myelopoiesis resulted in less resistance and more severe pathogen burden following to *Listeria monocytogenes* infection, rescued by the re-colonization of the mice gut with a complex microbial community (Khosravi et al. 2014). These findings reveal that gut microbiota directs innate immune cell development via promoting hematopoiesis. Conversely, the gut microbiota can also alter autoimmune conditions. GF mice develop significantly less severe diseases in models of experimental autoimmune encephalomyelitis (EAE) (Ochoa-Reparaz et al. 2010). On the other hand, colonization with segmented filamentous bacteria (SFB), promotes autoimmune arthritis through the induction of antigen specific T_H17 cells, which in turn promote auto-antibody production via B cell expansion in germinal centres (Wu et al. 2010). Additionally, commensal microbiota is also describe to be involved in inflammatory diseases, such as inflammatory bowel diseases (IBD), and in metabolic diseases (*e.g.* obesity, diabetes) (detailed in section 1.4.5).

All together, these evidences show that the contribution of the gut microbiota on the development and maintenance of the immune system goes far beyond the intestinal mucosa and shapes local and systemic immune responses to impact the overall physiology of the host.

The biological effect induced by gut microbiota are not exclusively mediated by PRRs, as evidenced by its huge impact on the overall host's physiology. Indeed a variety of cellular receptors and intracellular targets evolved to sense the myriad of microbial metabolites, intermediates and/or end-products of dietary constituents, which are released in the intestine.

Through the host's targets, microbial metabolites stimulate specific signalling pathways that ultimately modulate the host energy, nutritional and immune homeostasis and protect from pathogens. In the next section, the host-microbial crosstalk, mediated by cell receptors (*e.g.* G-protein coupled receptor (GPR), aryl hydrocarbon receptor (AhR)) and intracellular targets (*e.g.* histone deacetylase), will be detailed, and organized based on some interesting microbial metabolites.

1.4.4. Microbial-derived Metabolites in Host-Bacterial Cross-Talk

Whilst the microbiota phylogenetic composition may vary within healthy individuals, the functions performed by the gut microbial community are very similar. Indeed, a "functional core" has been identified to be essentially composed of genes encoding metabolic activities such as short-chain fatty acids (SCFA), essential amino acids (AA) and vitamin production in all individuals (Kurokawa et al. 2007). Additionally, gut microbiota produces indoles from tryptophan metabolism, trimethylamine (TMA) from choline degradation and metabolize toxins to less or more harmful substances contributing to the intestinal xenobiotic metabolism. Consequently, the co-habitation of bacteria and the host has been revealed to be mutually beneficial for both partners, in healthy conditions. Indeed, while the host provide ecological niches (*e.g.* mucus) and nutrients for the microbial community, bacterial colonization helps the host energy, nutritional and immune homeostasis as well as the protection against inflammation and pathogens, through the production of bacterial metabolites. Accordingly, comparative studies on germ-free (GF) and conventionally (CONV) raised mice revealed a correlation between the absence of intestinal microbiota with the reduction of intestinal vasculature, undeveloped gut-associated lymphoid tissue and alterations in nutrition and energy metabolism, all rescued upon the reintroduction of gut bacteria (Ha et al. 2014). These evidences further confirmed that the influence of bacterial metabolites is not localized in the colon but rather projected beyond the gut. Coherently, appetite regulation, gut mobility, energy balance and immune tone result from the integration of multiple signals from the intestinal ecosystem and bidirectional communication along the gut-brain axis (Holzer et al. 2012).

Microbial-derived SCFA

One evolutionary driving force for maintaining the intestinal microbiota is their role in energy yield, particularly through the fermentation of otherwise non-digestible compounds derived from the diet and present in the large intestine. The microbiota colonizing the large intestine, has access only to the dietary residues that escape digestion by the host's digestive enzymes along the digestive system. These residues are mainly insoluble plant fibers and resistant starch as well as different sorts of host derived glucans found in the intestinal mucus layer (Flint et al. 2008). With their large repertoire of glycosyl hydrolases, bacteria degrading plant-borne fibers play a unique ecological role by providing the rest of the microbial food chain with a variety of simple oligomers that serve as energy source for fermentative microorganisms. It is hence conceivable that diets with low fiber diversity can induce an underdevelopment of the corresponding hydrolytic microorganisms and thereby modulate the overall dominant microbiota.

From the fermentation of undigested carbohydrates, intestinal bacteria produced important signalling molecules such as **short chain fatty acids** (SCFA). SCFAs are organic acids consisting of 2–6 carbon atoms, among which the principal component in the intestine are acetate (C2), propionate (C3), and butyrate (C4). These fatty acids arise in the colon from the fermentation of carbohydrates, derived from dietary fibers or the inefficient digestion and absorption in small intestine. The relative ratio of acetate, propionate, and butyrate in human colon content is 57:22:21, respectively (Cummings et al. 1987), and the total concentration of these three SCFAs in the lumen is in the range of 50–150 mM (Cummings et al. 1987). The majority of SCFA production is locally utilized by the gut epithelial cells but significant amounts are also transported across the epithelium to distant tissues via the circulatory system.

Dependent on the luminal concentration, SCFA enter the colonic epithelium by diffusion or through effective transport systems in the apical membrane of IECs (*e.g.* proton-coupled and sodium-coupled monocarboxylate transporters - MCT), although, under physiological conditions, diffusion is supposed to be the major entry mechanism (Iwanaga et al. 2006). After absorption, the metabolic fate of the SCFA differs. Butyrate is metabolized primarily by the intestinal epithelium, where it is converted to ketone bodies or oxidized to CO_2 (Pennington 1952). Propionate is transported to the liver while acetate is taken up primarily by peripheral

tissues and both are used as substrates for energy metabolism and lipid synthesis (Wong et al. 2006). Around 6%-9% of the total energy intake for humans, accounts from SCFA absorption and up to 60-70% of the colonic energy need (Wong et al. 2006). However, the biological effects of SCFAs in the colon are not restricted to their role as energy substrates for the epithelial cells. Indeed, these microbial metabolites promote water and electrolyte absorption in the colon, thus providing protection against potential diarrheagenic diseases, modulate the mucosal immune system by helping the development of a tolerant environment and aid in the maintenance of the mucosal barrier (further detailed below).

Impact of SCFA on Host's Health

The biological effects of SCFA are mediated by both cell-surface receptors (G-protein coupled receptors) and intracellular targets (histone and lysine deacetylases as well as transcription factors).

Among the cell-surface receptors that mediate SCFA signalling, the **G-protein coupled receptors** (GPR) have been identified (*Figure 16*). GPRs are the largest and most diverse group of membrane receptors in eukaryotes, which transduce external stimuli. Extracellular ligand-binding to GPRs induces conformational changes that alter the receptor's interface with cytosolic effectors, thus generating a cascade-like signalling pathways. Butyrate, propionate and acetate are all agonists for GPR41 and GPR43, whereas for GPR109a only butyrate showed affinity (Le Poul et al. 2003; Thangaraju et al. 2009). Evidences from GPR-deficient mice identified a mechanism of SCFA in the modulation of energy balance, mainly via GPR41 and GPR43. Indeed, GPR41$^{-/-}$ and GPR43$^{-/-}$ mice revealed a reduced harvest of energy and a resistance to high fat diet (HDF)-induced obesity, insulin insensitivity, and dyslipidemia, respectively (Bjursell et al. 2011).

However, the impact of microbial-produced SCFA via GPR signalling did not relied only on energy metabolism. Gut hormones appeared to be as well influenced by SCFA through GPR signalling activation. For example, GPR41$^{-/-}$ mice showed reduced expression of peptide-YY (PYY), an enteroendocrine cell-derived hormone involved in gut motility and intestinal transit rate. Moreover, the production of propionate in the gut was shown to activate GPR41 in the nerve fibres of the portal vein, resulting in up-regulation of genes required in intestinal synthesis of glucose, or intestinal gluconeogenesis (IGN) (De Vadder et al. 2014).

The SCFA-induced GPR signalling appeared to play an important role also in modulating the immune system. Depletion of GPR43 in colitis, arthritis and asthma murine models, showed increased production of inflammatory mediators and enhanced recruitment of immune cells, attenuated by acetate supplementation (Maslowski et al. 2009). Additionally, SCFAs increase a GPR-dependent expression of the anti-inflammatory cytokine IL-10, with a consequent reduction of pro-inflammatory responses and proliferation of effector $CD4^+$ T cells (P. M. Smith et al. 2013).

GPR43 on colonic T cells has been described to regulate the size and function of the colonic Treg pool and protects against experimentally induced colitis in murine model (P. M. Smith et al. 2013). Coherently, GPR109a activation by butyrate was reported to stimulate the differentiation of Treg and IL-10 producers T cells and suppressed carcinogenesis (Singh et al. 2014). Taken together, these data suggest that SCFA-sensing GPRs play an important role in regulation of immunity and inflammation.

Beside the aforementioned surface receptors, the cell response induced by SCFA is mediated by intracellular targets like **histone and lysine deacetylases** (HDACs and KDACs, respectively), which modulate the epigenetics of target cells (*Figure 16*). It has been demonstrated that HDACi treatment induces significant changes in expression patterns of 5%–10% of human genes, and an even larger fraction of the transcriptome is affected, although most changes in expression are of low magnitude (Daly & Shirazi-Beechey 2006). While acetate is inactive, butyrate and propionate have been both described as HDACs inhibitors (HDACi). In particular, butyrate targets Class I and Class IIa HDACs (Schilderink et al. 2013) and influences the expression of 221 genes within the human genome, involved in proliferation, differentiation and apoptosis (Daly & Shirazi-Beechey 2006). Through the inhibition of HDAC, butyrate and propionate exerts multiple effects. For example they both stimulate IL-10 receptor α subunit (IL-10RA) in human IECs, inducing an epithelial barrier effect mediated by the repressed expression of the claudin-2-tight junction (Cldn2) protein, reported to increase in inflammatory bowel diseases (IBD) patients (Zheng et al. 2017).

Protein lysine acetylation has emerged as a key post-translational modification in cellular regulation, in particular through the modification of histones and nuclear transcription regulators. Microbiota has been suggested to have a role on protein lysine acetylation through the production of SCFA. Indeed, germ-free (GF) mice evidence an increase in lysine ε-

acetylation in colon and liver upon conventionalization, suggesting the role of microbiota on epigenetic regulation, likely through the production of SCFA (Simon et al. 2012).

Furthermore, various longer chain acylations of histones have been characterized, including crotonylation, butyrylation and hydroxybutyrylation (Tan et al. 2011; Chen et al. 2007; Xie et al. 2016). In particular histone crotonylation has been recently described to be abundant in the intestinal epithelium, especially the crypt fraction of the small intestine and the colon and been promoted by SCFA in intestinal cells and organoid cultures (Fellows et al. 2018). Coherently, the microbiota depletion through antibiotic treatment led to a loss of histone crotonylation, suggesting that the signalling between the microbiota and chromatin might be mediated through this post-translational modification (Fellows et al. 2018).

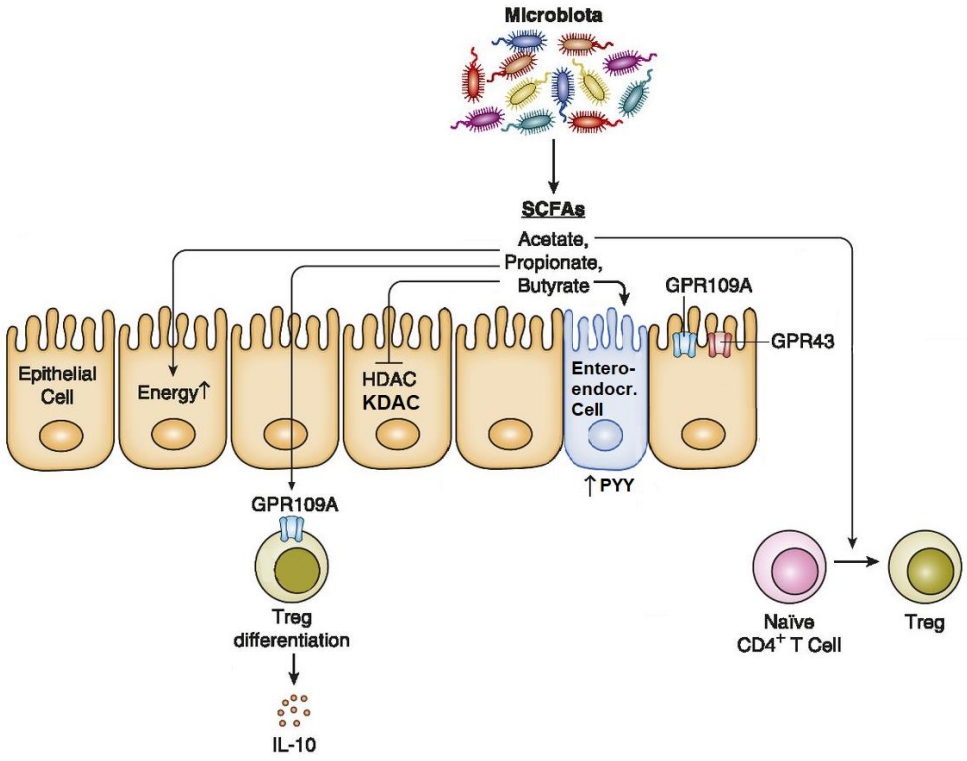


Figure 16: Microbiota-derived SCFA shape cellular metabolism in intestine. *Schematic representation of few mechanism for microbial-host interaction in intestine via SCFA (adapted from Blacher 2017).*

Recently, another mechanism for the microbial-derived SCFA to impact the host, has been evidenced. Exploring the mechanism involved in the regulation of angiopoietin-like 4 (ANGPTL4) in the human colon, Alex and co-workers identified a novel role of butyrate as ligand of the **transcription factor** peroxisome proliferator-activated receptor (PPAR) (S. Alex

et al. 2013). In this study they first reported that SCFA, and in particular butyrate, stimulate the synthesis of ANGPTL4, a regulator of peripheral lipid and glucose metabolism, whose expression is regulated by PPARs. The SCFA-induced stimulation of ANGPTL4 expression emerged to be mediated by transactivation and binding of butyrate to PPARγ, likely by serving as selective PPAR modulator. Additionally, structural modelling of butyrate on PPARγ shows butyrate bound in the ligand-binding pocket of the protein, coherently with other known agonists (*e.g.* decanoic acid) (S. Alex et al. 2013) (*Figure 17*).

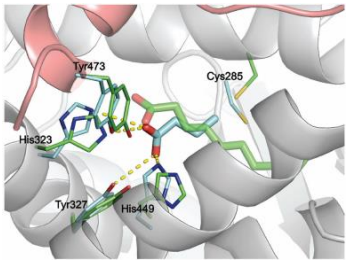


Figure 17: Modeling of butyrate into the PPAR binding pocket overlaid with the crystal structure of decanoic acid *(butyrate in cyan sticks, decanoic acid in green stick) (Alex 2012).*

This newly described role of butyrate as ligand of a transcription factor, shed light on the multiple targets of SCFA (cell-surface, intracellular targets as well as transcription factors) in eukaryotic cells. Additionally, from these reports as well as several other evidences in literature, it is clear that, through the production of SCFA, intestinal microbiota have multiple beneficial effects on various aspects of mammalian metabolism with a relevant role on human health (the role of microbial in pathogenesis is detailed in section 1.4.5).

Microbial metabolism of Tryptophan

The proteins reaching the colon (about 12-18g/d) are either residual from diet or endogenously produced (*e.g.* enzymes, mucins). The tryptophan (Trp) absorption in the intestine is primarily mediated by B^0AT1 (SLC6A19) transporter, expressed in different epithelial cells. In small intestine, B^0AT1 transporter is localized in the brush-border membrane of epithelial cells to guarantee the apical uptake of amino acids (Romeo et al. 2006; Gao et al. 2018). In addition to its primarily role as a substrate for protein synthesis, Trp is also metabolized in human through two metabolic pathways: the kynurenine pathway (KP) and the serotonin pathway. Approximately 95% of the ingested Trp is degraded to kynurenine, kynurenic acid, quinolinic acid, picolinic acid, and nicotinamide adenine dinucleotide (NAD) through KP, regulated by two rate-limited enzymes: tryptophan 2,3-dioxygenase (TDO) in the liver and indoleamine 2,3-dioxygenase (IDO) in extrahepatic tissues (Peters 1991). Approximately 1–2% of ingested Trp is converted to serotonin (5-HT) and melatonin via the serotonin pathway. However, gut microbiota can also use Trp directly, thus limiting the availability for the host. It was estimated that approximately 4–6% of Trp is metabolized by gut microbiota (Gao et al. 2018). Accordingly, the detected plasma concentration of tryptophan and *N*-acetyltryptophan in conventional (CONV) mice was 40-60% lower than in their respective germ-free (GF) counterparts (Wikoff et al. 2009).

The major tryptophan derivative arising from the microbial metabolism is the **indole**, produced by tryptophanase (TnaA), which can reversibly convert tryptophan into indole, pyruvate, and ammonia. Many Gram-positive and Gram-negative bacteria encode a single copy of the *tnaA* gene in their chromosome and produce indole. Even if several organisms contain the tryptophan biosynthesis pathway or express a TnaA homologue, to date, only those encoding *tnaA* can synthesize indole.

Indole has been detected in the mouse and human gut at concentrations around 250–1100μM, and identified in the blood, peripheral tissues, urine and brain at concentrations as high as 10-200μM. However, animals cannot synthesize indoles so these concentrations account only from the bacterial metabolism. Accordingly, various indole-containing molecules have been detected only in the plasma of CONV mice compared to GF, confirming the strict bacterial-dependent origin of these compounds. Consequently, the lower concentration of tryptophan

in plasma of CONV mice compared to GF, is likely resulted from the microbial tryptophanase activity (Wikoff et al. 2009).

Other indole-containing molecules are produced by the microbial metabolism of tryptophan, such as indole pyruvic acid, indole acetaldehyde, indole lactate and tryptamine, a neurotransmitter synthesized by gut bacteria, key in the enteric nervous system signalling (Lee 2015). Then, indole acetaldehyde can be converted to indole acetic acid and tryptophol, and the former in skatole. Moreover, indole lactate may be converted to indole acrylic acid and subsequently to indole propionic acid (Smith & Macfarlane 1997; Gao et al. 2018).

Within the complex ecological community residing in the gut, the microbial generation of tryptophan-derived metabolites is influenced by the quantity and quality of dietary fibers. In fact, as a primary carbon and energy source, fibers represents the major determinant of bacterial composition in the colon. Accordingly, it has been evidenced that sugar restriction in diet promotes the expansion of different *Lactobacillus* species that possess the enzymatic machinery to generate Trp metabolites (Zelante et al. 2013).

Mechanisms for the Biological effects of Indoles and Impact on Host's Health

In its unmodified form, indole serves as a signalling agent in bacterial communities, modulating spore formation by Gram-positive strain, plasmid stability, cell division, antibiotic tolerance, virulence in *Escherichia coli*, and biofilm formation in *Escherichia coli* and *Vibrio cholerae* (Lee et al. 2015).

Indole metabolites produced by intestinal microbiota, are targets for host's enzymes. For example, the hepatic transformation of indole, originates the indoxyl sulphate (indicant), a nephrotoxin that accumulates in the blood of patients suffering from chronic kidney failure (Tan et al. 2017) (*Figure 18*). Because the tryptophanase activity derives from only a subset of commensal bacteria, non-indole-producing bacteria, such as various *Bifidobacterium* species, have been administered as a test probiotic to dialysis patients, decreasing their plasma levels of indoxyl sulfate (I.-K. Wang et al. 2015). Conversely, a different set of commensal bacteria, in particular *Clostridium sporogenes*, perform the metabolic transformation of indole to indole-3-propionic acid (IPA), a powerful antioxidant currently investigated as a possible treatment for Alzheimer's disease (Bendheim et al. 2002). Plasma concentration of IPA, have been

identified only in CONV mice and recovered in GF mice only after the colonization with *Clostridium sporogenes*, confirming the strict microbial origin of this metabolite (Wikoff et al. 2009).

Recent works identify indole as a signalling molecule by which gut microbiota communicate with the host and influence its metabolism and intestinal homeostasis. In particular it has been evidenced that indole increases GLP-1 release during short exposures in L cells through two key molecular mechanisms (Chimerel et al. 2014). On the one hand, indole inhibited voltage-gated K^+ channels and led to enhanced Ca^{2+} entry, thereby stimulating GLP-1 secretion. On the other hand, indole slowed ATP production by blocking NADH dehydrogenase, thus leading to a prolonged reduction of GLP-1 secretion (Chimerel et al. 2014) (*Figure 19*).

Additionally, indoles have been described to targets several host receptors, including the **aryl hydrocarbon receptor** (AhR) (Hubbard, Murray, Bisson, et al. 2015) and **the pregnane X receptor** (PXR) (Venkatesh et al. 2014) (*Figure 18*). By ligand binding to PXR, indole-3-propionic acid (IPA) down-regulates the enterocyte mediated inflammatory cytokine tumor necrosis factor-α (TNF-α) with a concomitant up-regulation of junctional protein-coding mRNAs and up-regulation of TLR4, by a PXR-dependent mechanism (Venkatesh et al. 2014). Similarly, AhR activation in IECs protects mice from experimentally induced colitis by enhancing IECs barrier function via the increase of IL10 receptor expression and the enhancement of tight junction integrity (Liu et al. 2018; Yu et al. 2018). Coherently, GF mice show decreased level of tight junctions and adhesion junctions-associated molecules probably partly due to the lack of microbiota metabolism of Trp. Although the exact mechanism is not yet completely described, oral administration of indole stimulates the production of tight junctions and adhesion junctions-associated molecules as well as mucin production (Bansal et al. 2010), likely through the transcription factors PXR or AhR.

Additionally to the stimulation of tight junctions, AhR has also been evidenced to play a role in IL-22 production and consequents stimulation of antimicrobial peptides (AMPs) production and a protection against pathobionts by mechanisms likely to be dependent on ligand-binding (Yu et al. 2018; Zelante et al. 2013) (the role of AhR in physiology is detailed in section 1.5.4).

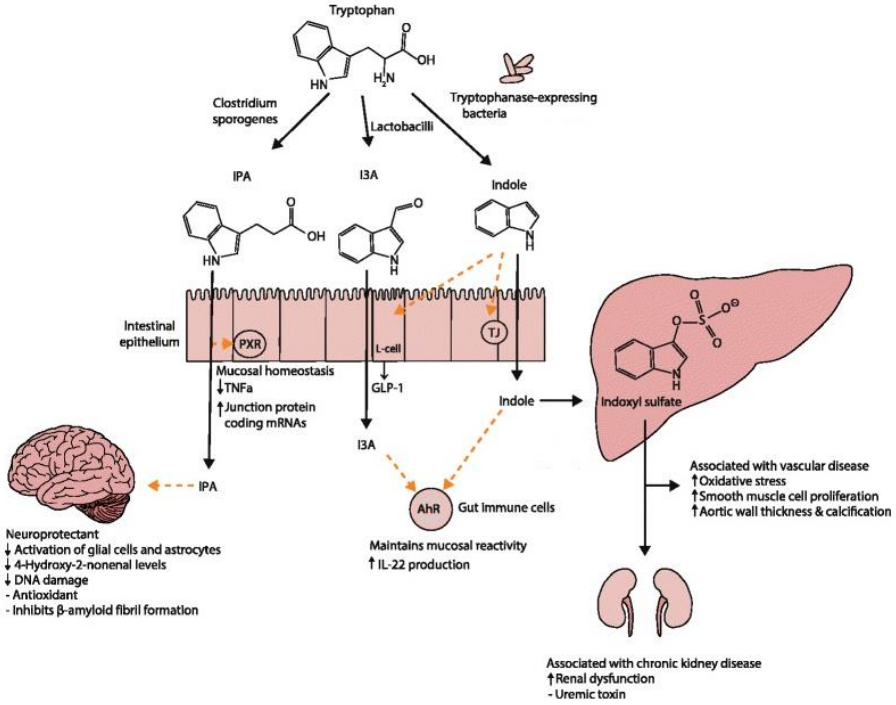


Figure 18: Synthetic molecular mechanism of action of indole and its metabolites in host physiology and diseases *(adapted from Zhang 2016).*

These evidences support the definition of indole-derivatives as beneficial interkingdom signal molecules that improve intestinal epithelial-cell function, maintain controlled inflammation, and additionally increase resistance to pathogen colonization.

Microbial metabolism of Choline and Cardiovascular Disease Risk

Beside the production of beneficial molecules, gut microbiota also produce potential harmful metabolites from nutrients, for example trimethylamine (TMA) from dietary choline. Choline is a dietary component, mainly present in eggs, red meat and wheat germ, in free and esterified form (such as phosphocholine, glycerophosphocholine, phosphatidylcholine, and sphingomyelin). It is crucial for normal functions of all cells and it was recognised as essential nutrient by the Institute of Medicine (IOM) in 1988 (Zeisel & Da Costa 2009). Indeed choline is the major dietary source of methyl groups, via the synthesis of S-adenosylmethionine, it is required for the synthesis of some phospholipids (phosphatidylcholine, lysophosphatidylcholine and sphingomyelin) and it also plays a role in supporting cholinergic neurotransmitters (Zeisel & Da Costa 2009).

The catabolism of choline, generates trimethylamine (TMA), acetate and ethanol as the final products of fermentation. The exclusive microbial catabolism of choline is provided, in human, by the commensal microbiota (Al-Waiz et al. 1992). TMA is a volatile tertiary aliphatic amine derived from the diet either directly, by consumption of foods rich in TMA, or indirectly, by the intake of TMA precursors such as choline and L-carnitine (Koeth et al. 2013) (*Figure 20*). When choline-rich food is ingested and the concentration of choline exceeds the capacity of its absorption in the small intestine, it lands in the large intestine where it is catabolized in TMA and dimethylamine (DMA) by commensal bacteria equipped with the choline utilization (cut) gene cluster (Craciun & Balskus 2012). *In silico* predictions suggested that several members of the human intestinal microbiota (including *Clostridium* spp., *Anaerococcus* spp., *Collinsella* spp., *Desulfitobacterium* spp., *Klebsiella* spp., *Escherichia* spp., *Providencia* spp., *Yokenella* spp. and *Proteus* spp.) have the ability to degrade choline to TMA (Craciun & Balskus 2012).

Once produced by microbial community, the physico-chemical properties of TMA allow its efficient absorption from the gastrointestinal tract with subsequent extensive *N*-oxidation by the hepatic flavin mono-oxygenase enzymes (FMO1 and FMO3), originating the trymethylamine *N*-oxide (TMAO) (Romano et al. 2015) (*Figure 19*). Although the physiological role of TMA in humans is still unclear, TMAO has been recognized as a pro-atherogenic metabolite (*Figure 19*).

In both retrospective and prospective studies, it was observed that plasma levels of TMAO and TMAO-precursors (including choline and L-carnititine), are associated with coronary artery diseases (CAD), cerebrovascular events and peripheral artery diseases (Wang et al. 2011; Koeth et al. 2013). Although the link between TMAO and atherosclerosis has been established, the exact mechanism through which TMAO enhances the cardiovascular disease risk is still elusive. The fundamental pathological changes associated with atherosclerosis is the formation of lipid-rich macrophages (also known as "foam cells") within the arterial wall. In this regard it was evidenced that TMAO facilitates foam cells formation in an atherosclerosis mice model (Apoe$^{-/-}$), when supplemented with choline or TMA. Conversely, the microbiota depletion of these mice through an antibiotic treatments, results in an inhibition of the TMAO-dependent foam cell formation (Wang et al. 2011). Furthermore, TMAO was demonstrated to suppress the reverse cholesterol transport (RCT), the process ensuring the efflux of cholesterol from peripheral cells and its transport back in the liver for its metabolism and biliary excretion (Koeth et al. 2013). However, the antibiotic treatment rescued the suppressed RCT (Koeth et al. 2013). Altogether these evidences indicate that TMAO accelerates atherosclerosis by facilitating cholesterol influx and inhibiting cholesterol efflux. Additionally, the rescued phenotypes observed upon antibiotic treatments, confirm the microbial-dependent origin of this proatherogenic metabolite.

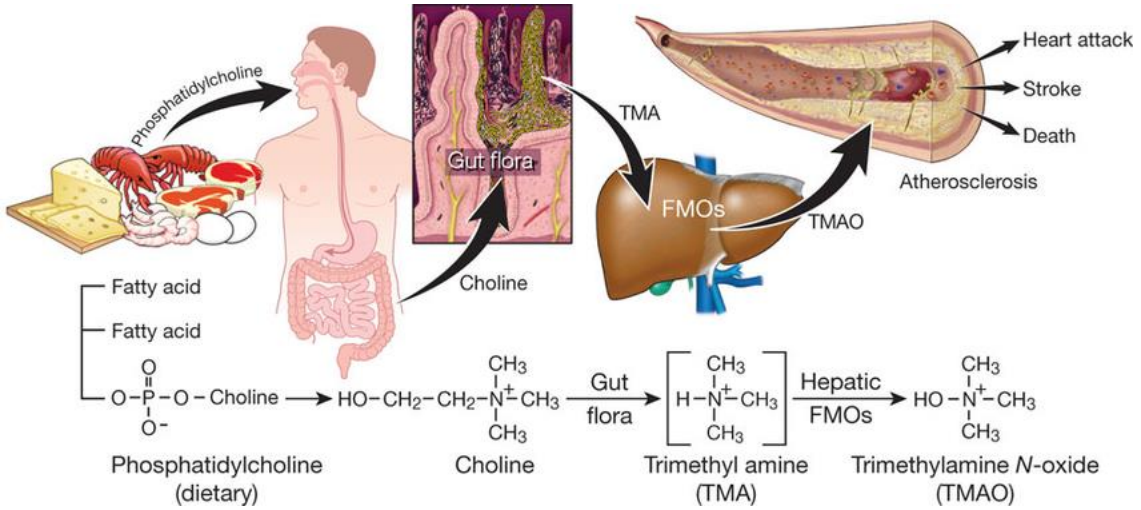


Figure 19: Gut microbiota-dependent metabolism of dietary choline and atherosclerosis.
Schematic summary for microbial pathway generating the pro-atherosclerotic TMAO (Wang 2011).

Microbial metabolism of Bile acids and Impact on Liver Health

An important class of bacterial metabolites that impacts host health are the secondary bile acids, major regulators of bacterial community and produced through the microbial metabolism of host's primary bile acids. Primary bile acids (cholic acid and chenodeoxycholic acid in human –CA and CDCA, respectively) are endogenous molecules synthesized from the cholesterol in the liver. Bile acids are released into the duodenum after a meal, to facilitate the absorption of triglycerides, cholesterol and lipid soluble vitamins. Bile acids and conjugated bile acids are then reabsorbed in the ileum and recirculated to the liver in a so-called enterohepatic circulation. Commensal bacteria, equipped with bile salt hydrolases (BSH), deconjugate the conjugated bile acids into secondary bile acids (deoxycholic acid, lithocholic acid, ursodeoxycholic acid – DCA, LCA and UDCA, respectively) to prevent active reuptake from the small intestine (Jones et al. 2008).

Bile acids appear to be a major regulators of the gut microbiota and the connection between liver health, faecal bile acids and microbial composition has been highlighted by different studies. Indeed, the progression of cirrhosis in patients was correlated with lower total, secondary, secondary/primary bile acid ratio and a coherent decrease in bacteria producers of secondary bile acid (*Lachonospiraceae*, *Ruminococcaceae* and *Blautia*) (Kakiyama et al. 2014).

Bile acids in the intestine have both direct antimicrobial effects on gut microbes, and indirect effects through the **farnesoid X receptor** (FXR)-induced antimicrobial peptides (Inagaki et al. 2006) (*Figure 20*). FXR is a bile acid-activated nuclear receptor that regulates the homeostasis of bile acids, lipids, and glucose and is activated by endogenous ligand including CA, CDCA, DCA, LCA and UDCA (Inagaki et al. 2006). In addition to FXR, the plasma membrane receptor **TGR5** is activated mainly by secondary bile acids LCA and DCA (Chen et al. 2011) and is therefore an interesting target in the context of microbiota-bile acid interactions.

Thus, through the production of bile acids and the modulation of enterohepatic circulation of bile salts, the host-microbiota impacts the liver health and regulates the ecological homeostasis in the gut, limiting potential pathogenic bacterial populations.

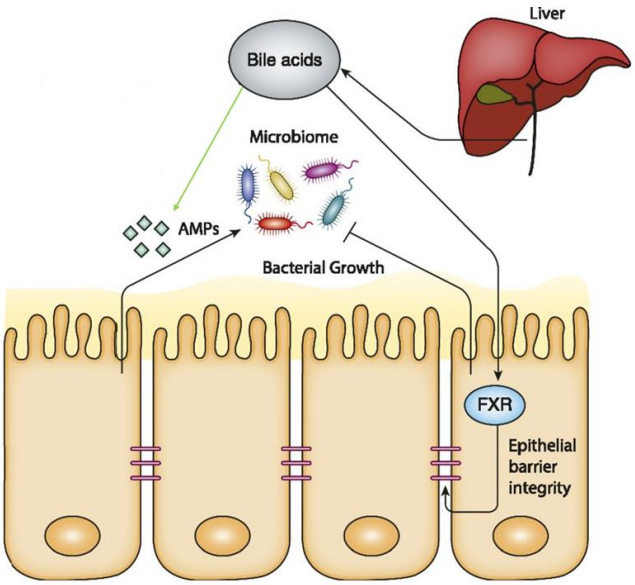


Figure 20: Bile acids in the intestine. *Schematic representation of few mechanism for microbial-host interaction in intestine via bile acids. Green arrow indicates the stimulation of AMP production. AMP: antimicrobial peptides, FXR: farnesoid X receptor (adapted from Blacher 2017)*

Beside these described functions of intestinal microbiota in gut ecology and homeostasis, the enormous potential of commensal bacterial in modulating host physiology is still poorly characterized.

By examining a metagenomic library from three patients for effectors activating NFκB, Brady's team identified 26 unique commensal bacteria effector genes, predicted to encode proteins with diverse catabolic, anabolic and ligand-binding functions. Among these, one effector gene family recovered from all the three patients, encodes for the production of a *N*-acyl-3-hydroxy palmitoyl glycine (commendamide). This metabolites, identified in cultures of *Bacteroides vulgaturs*, activates the human GPCR (GPR132/G2A) *in vitro* with a potentially important implications for autoimmune disease and atherosclerosis (Cohen et al. 2015). These newly described microbial family of molecules, point out to the enormous potential of commensal bacteria for modulating the host's physiology and the needed for expanded knowledge regarding the produced metabolites as well as characterization of their physiological role.

1.4.5. Dysbiosis in Human Pathology

As evidenced in the previous sections, the host-micrbiota cross-talk is based on a complex equilibrium to maintain the physiological and ecological homeostasis. Thus it is not surprisingly that distortion in the gut ecology could lead to potential negative effects for the host's health. An adverse alteration of the microbiota composition, termed dysbiosis, is indeed correlated with different human pathological states. However, rather than a simple microbial change in composition, the dysbiosis has to be considered as an alteration of the complex man-microbes symbiosis influencing both ecology and physiology in the intestine. As such, dysbiosis contributes to the underlying pathophysiology of a wide range of diseases, including obesity, diabetes, inflammatory bowel diseases, non-alcoholic fatty liver diseases and cardiovascular diseases (Aron-Wisnewsky & Clément 2016).

Several defining features of dysbiosis have emerged through the investigation of this patho-associated microbial population: a reduced microbial diversity and a decrease in organisms, considered beneficial to human health, with a corresponding increase in pathobionts, members of normal commensal microbiota with the potentiality to cause pathology (Nibali & Henderson 2016). On taxonomic level, dysbiosis is generally characterized by a depletion of obligate anaerobic bacteria such as *Bacteroides* spp. and *Ruminococcus* spp., and conversely an increase in facultative anaerobes, including the family of *Enterobacteriaceae* (*Escherichia coli*, *Klebsiella* spp., *Proteus* spp.), as reported in different pathological states such as Crohn's diseases, ulcerative colitis, *Clostridium difficile* infections and obesity (Manichanh et al. 2006; Lepage et al. 2011; Le Chatelier et al. 2013). These changes in microbial abundance and composition is traduced in a major impact on microbial derived metabolites, such as a decline in SCFA production. However, microbiota assessments in case–control studies can only reveal associations, and cannot directly establish the causal contribution of the dysbiosis on the pathogenesis. Nevertheless, delineating the causes and etiological consequences of disease-associated dysbiosis remains a crucial challenge in studies of the human microbiota.

Among the plethora of human pathologies correlated with dysbiotic states, the herein discussion will be restricted to the metabolic diseases and the inflammatory bowel diseases (IBD).

Dysbiosis in Metabolic Diseases

Metabolic diseases are becoming a world-wide epidemic, with a dramatic increase in the prevalence of obesity and closely related diseases, such as metabolic syndrome, type 2 diabetes mellitus (T2DM), cardio vascular disease, and chronic kidney disease. Changes in eating habits, with an increase fat intake, are involved in the increased occurrence of metabolic diseases, such as obesity and diabetes, which are bearing features of the metabolic syndrome.

Although treatments are available to manage some of those diseases, these are not curative, evidencing a critical need for identifying novel potential targets and pathways involved in these diseases.

The major metabolic consequence of a high-fat diet is the insulin resistance associated with low-grade chronic systemic inflammation and an altered microbiota composition (Wellen & Hotamisligil 2005; Ley et al. 2005). The relationship between gut microbiota and metabolic diseases was first reported by Gordon's team, by demonstrating that leptin deficient mice, contained fewer *Bacteroidetes* and more *Firmicutes* than control mice (Ley et al. 2005). This study provided the first direct evidence of differences in the microbiota composition between in lean and obese animals. In support of that, a large reference catalogue of gut bacterial genes generated from high-throughput sequencing data, has evidenced the association between dysbiosis and cardiometabolic diseases (Li et al. 2014; Thomas & Ockhuizen 2012). Coherently, the colonization of GF mice with the faecal content of their CONV counterparts, induce a shift from a resitance to high fat diet (HFD)-induced obesity to an obese phenotype, without any change in calories ingestions (Bäckhed et al. 2004). Additionally a study conducted on twin pairs discordant for obesity revealed that the obese phenotype could be transferred to recipient mice whose will then show metabolic alterations and gained more body weight, compared to animal receiving the faecal microbiota from lean donors (Ridaura et al. 2013). These findings demonstrated firstly, the relevance of gut microbiota in these pathogenesis and secondly, that the gut microbial ecosystem is transmissible, at least in mice, opening the way to a therapeutic application of faecal transplantation. However, the implication of diet is still relevant in both the microbial modulation and the consequent development of the pathologic state.

HFD led to an increase in the *Escherichia coli* co-localization with DC in the intestinal lamina propria. Mice fed with a HFD display a systemic level of *E. coli* rising until a diabetic status is established, with a marked presence of bacterial DNA in various tissues (Amar et al. 2011).

This infection targets the mesenteric adipose tissue through a mechanism described as bacterial translocation, which requires the recognition of MAMPs by CD14, signalling through MyD88 and is leptin-regulated (Amar et al. 2011). Treating these animals with a probiotic strain *Bifidobacterium animalis* (sp. lactis 420), results in a decrease bacterial translocation and improved insulin sensitivity suggesting a potential therapeutic role of probiotics in the treatment of adverse metabolic phenotypes (Amar et al. 2011). Coherently, *Akkermansia muciniphila*, a mucine-degrading Gram-negative bacteria producer of butyrate, added to a HFD, protects mice from the development of insulin resistance and reduced tissue inflammation with a consequent improvement in gut integrity (Everard et al. 2013).

Moreover, studies performed on French and Danish cohorts, evidenced a reduced ecological gene richness correlated with higher inflammation tone, adiposity and insulin resistance in human (Cotillard et al. 2013; Le Chatelier et al. 2013). Additionally obese individuals among the lower bacterial richness group also gain more weight over time (Le Chatelier et al. 2013).

46 genera differed between the low gene count (LGC) and high gene count (HGC) groups, with higher abundance of Proteobacteria and Bacteroidetes in low gene count individuals and more Verrucomicrobia (*Akkermansia muciniphila*), Actinobacteria, and Euryarchaeota identified in the HGC group (Le Chatelier et al. 2013). Notably, low gene richness has also been associated with under-representation of *Faecalibacterium prausnitzii*, a butyrate-producing bacteria extensively described for its anti-infammatory effects (Sokol et al. 2008). Additionally, the reduction in *Faecalibacterium prausnitzii* abundance has also been reported in patients with T2DM (Hippe et al. 2016).

Overall, these described observations open diverse possibilities to identify new targets involved in the development of metabolic syndromes as well as leading to a personalized therapeutic approaches.

Dysbiosis in IBD

The incidence and prevalence of IBD is increasing in industrialized countries and, affecting generally young population, has a consequent huge socio-economic impact (M'Koma 2013). Like for metabolic diseases, treatments are available to dampen the immune response and the mucosal damage, but they are not curative. Crohn's disease (CD) and ulcerative colitis (UC) are both major categories of IBD characterized by inflammation of the GI tract, affecting any part from the mouth to the anus in CD whilst limited to the colon in UC.

IBD has long been recognized to have a genetic basis, indeed, studies conducted on monozygotic twins revealed a phenotypic concordance in 50–75% of CD patients, and the relative risk of developing CD is 800-fold greater compared to the general population (Halme et al. 2006). However in UC patient, only limited concordance has been revealed (10–20%), suggesting that heritability is less important in this pathology (Halme et al. 2006). The genetic predisposition in IBD has been further established by large genome-wide association studies, which have implicated a total of 163 IBD loci involved in innate immunity and host-microbes interactive pathways (Jostins et al. 2012). Surprisingly, it has been evidenced that IBD risk loci vary remarkably between different populations, *e.g.* NOD2 and autophagy genes are the major risk loci in Caucasian population, absent in Asian individuals (Thia et al. 2008). Although, most of the individuals with genetic susceptibility do not develop IBD, suggesting that the genetic component is not enough for IBD pathogenesis and that the environmental factors have also a major role. Accordingly, IBD development was observed in migratory group to high prevalence countries as well as in citizen from countries in rapid Westernization (Kaser et al. 2010).

The role of environmental factors such as life-style and, in particular, gut microbiota, have been reported in numerous studies. A first association of IBD with the presence of adherent-invasive *Escherichia coli* (AIEC) strain LF82 in ileal mucosa (Darfeuille-Michaud 2002), led to the hypothesis that a single pathobiont could be responsible for IBD pathogenesis. However, expanded knowledge on microbiota regarding IBD, point to a more general alteration of host-bacterial crosstalk rather than a single bacterial strain. Thus, conversely to the pathobiont hypothesis, dysbiosis has been evidenced as pivotal in the development of IBD. At taxonomical level, a decreased diversity is observed in CD and UC patients (Manichanh et al. 2006; Lepage et al. 2011) although the consequences are still unclear, particularly whether they

are cause or consequence. Nevertheless, the ecological changes observed in IBD patient and mimicked by experimental-induced colitis, are generally characterized by the reduction in strict anaerobes, particular Firmicutes, and an increase in Proteobacteria.

A specific member in Firmicutes, *Faecalibacterium prausnitzii*, was significant reduced in microbiota from CD patient and it was associated with a higher risk of post-operative recurrence of the ileal CD (Sokol et al. 2008). Moreover, in experimental-induced colitis, *Faecalibacterium prausnitzii*, exhibits anti-inflammatory effects likely due to secreted metabolites able to block NFκB activation and IL-8 production (Sokol et al. 2008). Beside *Faecalibacterium prausnitzii*, *Clostridium leptum* phylogenetic group was also identified as markedly reduced in CD patients (Manichanh et al. 2006). Considering the high efficiency of both *Faecalibacterium prausnitzii* and *Clostridium leptum* in SCFA production, it is conceivable that microbial metabolites with anti-inflammatory activity are depleted in IBD patients. Coherently, butyrate-producing bacteria, reduced in IBD patients, have been evidenced to enhance intestinal epithelial barrier integrity and thus proposed as a potential therapeutic tool for CD treatment (Geirnaert et al. 2017).

Beside SCFA production, other microbial metabolites have been identified as beneficial in IBD conditions. In particular tryptophan (Trp)- and indole-derivatives have been described for their protective role in IBD, mediated by **Aryl hydrocarbon Receptor** (AhR)-signalling pathway (discussed in detail in the specific section). AhR activation by ligand administration and AhR knock-down in mouse, suppressed and enhanced DSS-induced colitis respectively, suggesting a beneficial effect of AhR activation in IBD (Takamura et al. 2011; Furumatsu et al. 2011). Alterations in the composition of the intestinal microbiota as well as genetic polymorphisms detected in IBD patients, have been shown to impair the generation of the protective AHR ligands, ultimately contributing to immune dysregulation and disease pathology. In particular IBD-associated single nucleotide polymorphism (SNP) within the *CARD9* gene, affects microbiota composition altering the production of bacterial AhR ligands consequently with an increase in intestinal inflammation (Lamas et al. 2016).

Additionally, the activation of another nuclear receptor, **PPAR**, has been described as protective in experimental-induced colitis with a markedly improvement in macroscopic and histologic scores, a decrease in TNF α and IL-1β mRNA levels, and reduction of NFκB DNA binding activity (Desreumaux et al. 2001). Considering the recent description of butyrate as

ligand for PPARγ receptor (Alex 2013), it is likely that the described protection in colitis is partially mediated by microbial metabolites activating PPAR signalling. Thus, the administration of probiotics strains able to produced anti-inflammatory molecules, could have beneficial effects on IBD patients in term of slowing or preventing the progression of the disease process.

The evidence for a strong contribution of the gut microbiota to the onset of host physiology and pathology is growing. The use of germ-free rodent models has enabled to establish the molecular basis of the interactions between gut microbes and the physiology of the host. The modifications in the gut microbial ecology by environmental factors (*e.g.* dietary habits, antibiotics) that were observed in rodents and humans have further evidenced the modulatory roles of the gut microbiota and its contribution to host diseases. Additionally, mechanistic studied highlighted key microbial metabolites that modulate intestinal immune homeostasis and barrier protection by targeting different host receptors or intracellular targets. Furthermore, the role of microbiota has been evidenced beyond the gut, with stimulated signalling pathways able to influence cell population and metabolism far distant from the gut. Additional investigations are thus essential to improve the knowledge on the mechanisms and the physiological consequences for this critical host-microbiota crosstalk.

1.5. *Aryl Hydrocarbon Receptor: Description, Characterization and Physiological Role*

The environment offers vital benefits (food and energy), as well as deadly stimuli like infections and toxins. Both of these kinds of stimuli challenge the organism and, in particular, the barrier organs, which have to discriminate between them. Like the immune system, animals have evolved a variety of surveillance mechanisms in order to recognize foreign chemicals and trigger a specific response. One fundamental component of this surveillance consists of sensory proteins, as receptors, coupled to signalling molecules that initiate a biological response, often resulting in induction of enzymes to facilitate the biotransformation and elimination of the compound. In the early '70s, with the increasing interest in toxicology and environmental pollution, a ligand-activated transcription factor called aryl hydrocarbon receptor (AhR) has been firstly identified as a modulator of the toxic response to dioxins. Moreover, recent evidences have pointed to an equally important role of AhR as endobiotic receptor, which responds to a wide array of endogenous chemicals (endobiotics) and to metabolites produced by commensal microbiota. Additionally, the abnormalities and defects observed in AhR deficient mice, evidenced the role of this transcription factor in the overall physiology. Thus, this broadened knowledge about AhR, encouraged researchers to deeply investigate the signalling mechanism, as well as the role of this receptor beyond the mere toxicological effects, such as its impact on cell metabolism and immune system.

1.5.1. Expression and Localization

AhR is a ligand-activated transcription factor, with homologs identified throughout the animal kingdom. AhR homologues have been identified in **invertebrates**, such as in the nematode *Caenorhabditis elegans* and in the fruit fly *Drosophila melanogaster* (Huang et al. 2004; Wernet et al. 2006). Additionally, AhR is expressed in **vertebrates**, such as bird, amphibians and fish. In fish, two AhR forms, AhR1 and AhR2 have been described, and shear some features with the human AhR (Abnet et al. 1999). Regarding the mammals, AhR has been extensively studied in mice and human.

In the early '60s AhR was formerly identified in some inbreed **mouse** strains and characterized to be inducible by polycyclic aromatic hydrocarbons (PAHs) (Nebert et al. 1969). Few years later, Poland's group showed that an halogenated aromatic hydrocarbon (HAH; 2,3,7,8-Tetrachlorodibenzodioxin –TCDD) is a potent AhR inducer and overcome a "non-responsiveness" to a PAH (3-methylchlorantrene- 3-MC) in a number of mouse strains, indicating the existence of "responsive" and "non-responsive" mice (Poland et al. 1974) toward AhR activator and, more precisely, AhR ligands. The observation of responsive and non responsive mouse strains encouraged the researches to investigate the molecular basis of this polymorphism. The characterization of C57BL/6 mouse as responder (or responsive) strain conversely to DBA/2J as non-responder (or non-responsive) brought to the identification of $mAhR^b$ (from C57BL/6) and $mAhR^d$ (from DBA/2J) alleles responsible for high-affinity and low-affinity to AhR ligand, respectively (Poland & Glover 1990). A mutation from Ala375 to Val and a single point mutation in two specific protein regions (ligand-binding domain –LBD and PAS domain, respectively) has been identified as responsible for a reduced ligand affinity in DBA/2J mice (Ema et al. 1994).

In **human**, AhR was identified in the 70s as strongly inducible by the AhR ligand TCDD (Poland & Glover 1973). Comparing the human AhR allele (hAhR) with the high affinity $mAhR^b$ and the low affinity $mAhR^d$, it was evidenced that hAHR structurally resemble $mAhR^d$ rather than the $mAhR^b$, although an elongated C-terminal sequence and a Val381 substitutions (instead of the mouse Val375) differentiate the human receptor (Harper et al. 1988; Ema et al. 1994). Coherently human AhR showed a 10-fold lower affinity for TCDD compared to C57BL/6 mice allele.

As AhR is ubiquitously expressed in mammals, increasing interests have been focused on its expression at organ and tissue level, to highlight its physiological roles. It emerged that AhR expression differs significantly among tissues: it is not or weakly expressed in muscle tissue, testis, kidney and brain, while it is highly expressed in liver and in barrier tissues such as skin, gut, lung, mucosal epithelia and placenta, coherently with its role of sensor receptor of environmental factors (Frericks et al. 2007). AhR is expressed as well by different intestinal immune cells such as IELs, ILCs, macrophages, DCs and neutrophils (Frericks et al. 2007). Low AhR expression was evidenced in naïve T cells, helper T cells T_H1 and T_H2, and regulatory T cells (T_{reg}); conversely, high expression was detected in T_H17 cells and in IL-17/22-producing and -non producing subsets of peripheral γδT cells (Veldhoen et al. 2008). Additionally low AhR expression has been reported in naïve B cells from the spleen (Marcus 1998). In natural killer (NK) cells, AhR is moderately expressed and its activation was shown to stimulate antitumor activities as well as resistance to infections (Shin et al. 2013).

Considering the ubiquitous distribution of the receptor it is not surprisingly that its deletion generates significant defects such as vascular abnormalities, skin lesions and defect in reproduction and immune system in mice. Thus, the detailed study of the AhR activation pathway, its ligands and its regulation appeared to be imperative to identify its role in physiology and pathology.

1.5.2. Structure and Functional Domains

Ligand-activated transcription factors are a class of proteins able to rapidly respond to environmental signals and induce the expression of specific target genes. A member of this class is the basic helix-loop-helix (bHLH) protein superfamily, found in organisms from yeast to human and whose functions are critical in xenobiotic metabolism and developmental processes, including sex determination and the development of the nervous system and muscles (Jones 2004).

The bHLH protein family is structurally defined by two α-helixes linked by a loop, which provides flexibility. This family is further divided in different classes and groups according to their tissue distribution, DNA-binding specificities, dimerization potential and evolutionary relationships. More recently, other groups of this family have been identified by the presence of characteristic domains, among which the PER-ARNT-SIM (PAS) motif, defining the so-called bHLH-PAS family group, that includes AhR (Jones 2004; Fribourgh & Partch 2017).

Despite differences in their target genes and modes of regulation, bHLH-PAS transcription factor share a common domain architecture, consisting of bHLH DNA-binding domain followed by tandem PAS domains and a *C*-terminal region including a transactivation domain (TAD) (Fribourgh & Partch 2017) (*Figure 21*).

Similarly to other nuclear receptors, bHLH-PAS transcription factors require the dimerization with other protein partners in order to recognize a specific DNA sequence called E-box and consequently exert their transcriptional functions (Jones 2004). The formation of an heterodimer occurs within Class I, usually tissue restricted or signal specific, and Class II subunits, ubiquitously expressed and activated by a variety of stimuli (Fribourgh & Partch 2017). The dimerization is assured by PAS regions, among which PAS-A mainly prevent improper interaction with non-PAS containing molecules. The so-formed heterodimer is thus able to recognize and bind the specific E-box through bHLH motif as well as non-classical E-box sequences (Pongratz et al. 1998). An additional sequence in the PAS region, PAS-B, is involved in binding chaperone proteins and mediate the degradation of the receptor (Fribourgh & Partch 2017).

***Figure 21: Schematic representation of the bHLH-PAS protein domains**. In green the basil helix-loop-helix (bHLH) domain; in blue the two PAS domains (PAS-A and PAS-B); in red the transactivation domain (TAD) and in grey the variable regions.*

Examples of class I bHLH-PAS proteins include AhR (regulated by xenobiotics, endogenous and microbial-derived ligands), hypoxia-inducible factor-α (HIF-α, regulated by hypoxia), neuronal PAS domain proteins (NPAS, developmentally regulated), and circadian locomotor output cycles protein kaput (CLOCK, circadian rhythms), while a predominant member of Class II is the AhR Nuclear Translocator (ARNT) (Fribourgh & Partch 2017).

Within the bHLH-PAS family members, AhR has a special role as sensor of chemically diverse xenobiotics, environmental stresses and endogenous ligands, as well as microbial-derived ligands. Structurally related to other bHLH-PAS proteins, AhR has some distinct characteristics that make it the only mammalian bHLH-PAS transcription factor known to be regulated by binding to endogenous ligands (Fribourgh & Partch 2017). Endogenous and exogenous molecules bind to a specific ligand-binding domain (LBD), residing in the PAS-B region (Fribourgh & Partch 2017). Additionally, in the PAS-B domain, the interaction with chaperon proteins, such as the heat-shock protein 90 (HSP90), is established, and further stabilized by a second HSP90 site, residing within the bHLH region (Perdew & Bradfield 1996) (*Figure 22*). The *N*-terminus site of AhR, in addition to the bHLH motif, harbours two sequences commonly found in proteins shuttling between the cytoplasm and the nucleus: a Nuclear Localization Signal (NLS) and a Nuclear Export Signal (NES), both residing in the bHLH motif (Ikuta et al. 1998) (*Figure 22*). Additionally, a second NES has been identified in the PAS-A repeat and it seems to be mainly involved in nuclear export of unliganded receptor (Ikuta et al. 1998) (*Figure 22*). Similarly to other bHLH-PAS transcription factor, the DNA-binding site resides in the bHLH motif, recognizing a specific Xenobiotic Response Element (XRE, or DRE for Dioxin Respose Element) (*Figure 22*).

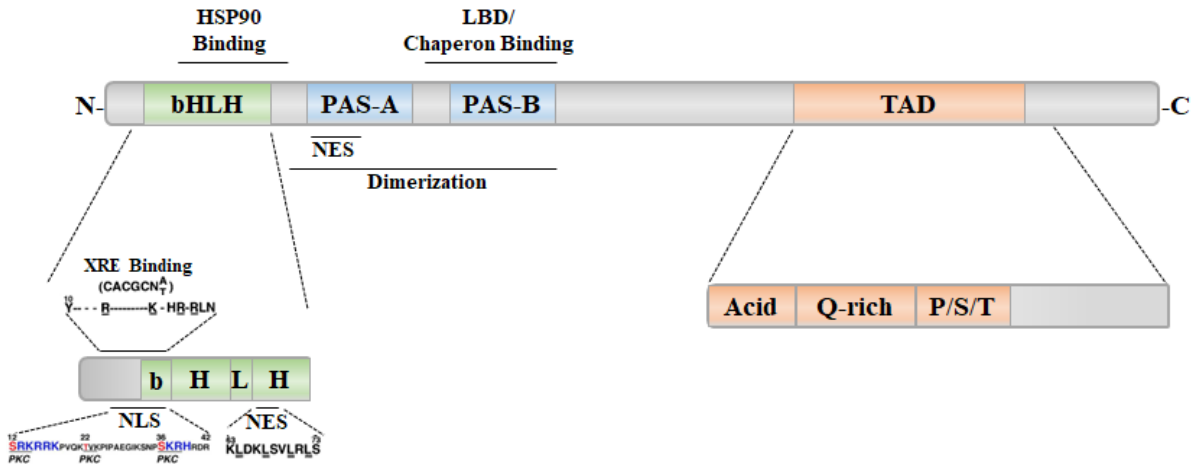


Figure 22: Schematic representation of AHR domains architecture. *In green the basic helix-loop-helix (bHLH) domain with the XRE recognition site, nuclear localization sequence (NLS), nuclear export sequence (NES) and them core sequence with putative consensus PKC phosphorylation sites in NLS (Ikuta et al. 1998); in blue the two PAS domains (PAS-A and PAS-B) with a second NES region in PAS-A; in orange the transactivation domain (TAD) composed by acidic subdomain (Acid), glutamine-rich subdomain (Q-rich) and proline/serine/threonine-rich subdomain (P/S/T).*

To exert its role at transcriptional level, AhR forms a heterodimer with its Class II partner ARNT, thus recognizing the XRE element in the promoter of downstream genes. The formation of the AhR:ARNT heterodimer is mediated by PAS-B domain provide a secondary dimerization surface in addition to bHLH region, while PAS-A stabilizes the dimer. Crystal structure of the mammalian AHR:ARNT heterodimer in complex with the XRE reveals that ARNT curls around AHR into a highly intertwined asymmetric architecture, with extensive heterodimerization interfaces and inter-domain interactions. The specific recognition of the XRE motif is thus locally determined by the DNA-binding residues and is globally affected by the dimerization interfaces and inter-domain interactions (Seok et al. 2017) (*Figure 23*).

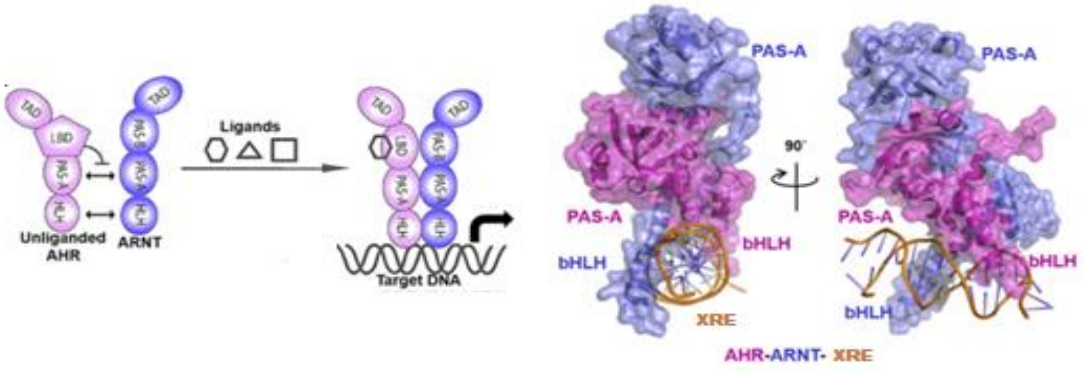


Figure 23: Overall architecture of AHR:ARNT heterodimer in complex with XRE (Seok et al. 2017). *Left: Schematic illustration of domain arrangement of AHR (magenta) and ARNT (blue) and assembly of the transcriptional complex; Right: Overall structure of AHR:ARNT:XRE complex in two perpendicular views.*

Both the AhR and its partner protein ARNT contain a transactivation domain (TAD) in their *C*-terminal region, mediating the transcription initiation by recruiting transcription factors and co-regulators to the transcriptional site. Detailed analysis revealed that the TAD consists of three subdomains, each of which has a low transactivation potential itself, but is able to act in synergy with the others. The TAD of the human AhR consists of an acidic-, glutamine (Q)-, and proline/serine/threonine (PST)-rich subdomains, each of which has critical roles: the Q-rich is important for the transactivation potential, the P/S/T-rich enhances transcriptional activity and the acidic one acts in synergy with the other two to induce activation of a XRE-driven reporter gene (Kumar et al. 2001) (*Figure 22*).

Additionally, receptor variants truncated in *C*-terminal region and TAD are improper degraded upon ligand activation, suggesting a role of TAD in the proteasomal degradation of AhR (Ma & Baldwin 2000) (proteosomal degradation further detailed in the specific section).

1.5.3. AHR pathway

Signalling pathway

In its inactive state, AhR is present in the cytoplasm and exists as a multi-protein complex with two chaperon proteins heat-shock protein (HSP) 90, co-chaperon p23 and AHR associated protein 9 (ARA9, also known as XAP2 or AIP) (Meyer & Perdew 1999; Meyer et al. 1998). The interaction of HSP90 with AhR is ensure by two chaperon-binding sites residing in PAS-B and bHLH domains (Antonsson et al. 1995; Perdew & Bradfield 1996) (refer to the section 1.5.2). Numerous studies evidenced that HSP90 is required for the proper folding and stability of the AhR complex. Indeed, cells treated with geldanamycin, a HSP90 inhibitor, exhibit a rapid proteolytic turnover of AhR (Chen et al. 1997). Coherently, reduced HSP90 expression levels in yeast, results in the arrest of AhR signalling (Carver et al. 1994), demonstrating the relevance of this chaperon protein for a functional AhR. Additionally, the dissociation of hsp90 *in vitro* leads to an inability of AhR to bind its prototypical ligand TCDD, further suggesting that hsp90 maintains the AhR in a conformation required for ligand binding in mouse and human (Pongratz et al. 1992; Coumailleau et al. 1995). However, the stability of the AhR:HSP90 complex was demonstrated to be higher in mouse, whilst the human complex needs molybdate to ensure its stability (Manchester et al. 1987). Altogether these evidences contribute to the description of hsp90 as critical for the proper ligand binding conformation and the activation pathway of AhR after its initial synthesis.

In addition to a dimer of hsp90, the presence of the co-chaperon p23 was identified in the AhR cytoplasmic complex, associated to the chaperon dimer. In yeasts, the inhibition of HSP90 ATPase by p23 was evidenced in presence of mutated hsp90 (Cox et al. 2004) and results in the stabilization of the HSP:AhR complex and the ligand responsiveness of the receptor (Kazlauskas et al. 1999).

Interestingly, the human hepatitis B virus X-associated protein (XAP) 2 has also been identified as part of the AhR cytoplasmatic complex, interacting with both hsp90 and AHR, via the PAS domains (Meyer et al. 1998; Meyer & Perdew 1999), and involved in stabilizing the cytoplasmatic AhR:HSP90 complex and protecting it from ubiquitination (Lees et al. 2003). The presence of XAP2 in the cytoplasmic AhR complex was supported by two independent studies in which the murine homolog of XAP2 (AhR-associated protein, AIP) (Ma & Whitlock

1997) and ARA9 (a protein sharing 99.9% of identity with XAP2) (Carver & Bradfield 1997) were identified bound to AhR:HSP90 complex.

Hence, the chaperon complex maintains AHR in an inactive and stable conformation with a high-binding affinity for ligand and retains the receptor in the cytoplasm, by masking its nuclear localisation signal (NLS) (Ikuta et al. 1998). Upon ligand-binding, AhR undergoes conformational changes leading to the dissociation of p23 and XAP2, the unmasking of NLS and the consequent translocation in the nucleus through the interaction with importin β (Ikuta et al. 2000; Petrulis et al. 2003). The ligand-dependent nuclear import of AhR is negatively regulated by phosphorylation for Ser-12 or Ser-36 at the two phosphor kinase C (PKC) sites adjacent to the bipartite NLS of AhR (refer to section 1.5.2, *Figure 22*) (Ikuta et al. 2004), suggesting a two-step mechanism in the ligand-dependent nuclear translocation of AhR involving firstly a phosphorylation event, then the binding to importin β for the nuclear shuttling of the protein. The nuclear translocation of AhR was originally hypothesized as mediated by the Class II partner protein AhR Nuclear Translocators (ARNT). However, in ARNT-deficient Hepa1 cells, the TCDD-induced nuclear translocation of AhR was observed (Pollenz et al. 1994), evidencing that ARNT is not directly involved in the nuclear localization of the receptor. Additionally, the intracellular trafficking of AhR has been described as influenced by the Q-rich and PST-rich regions of the TAD domain, by regulating both nucleoplasmatic shuttling and receptor activation (Tkachenko et al. 2016).

Upon nuclear translocation, HSP90 is released exposing HLH and PAS-A domains, essentials for the formation of a AhR:ARNT heterodimer (Heid et al. 2000; Chapman-Smith et al. 2004; Hao et al. 2011). Then, the AhR:ARNT complex binds to a consensus sequence TNGCGTG called xenobiotic response element (XRE) or drug responsive element (DRE) (Patel 2008). In contrast to the majority of bHLH transcription factors which bind the E-box sequences CANNTG, the XRE sequence is not symmetrical. Indeed, protein-DNA crosslink studies have shown that AHR and ARNT recognise 5'-TnGC and GTG-3' XRE half sites respectively (Swanson et al. 2002; Bacsi et al. 1995). This could explain why, the dissociated hsp90-form of AhR, in absence of ARNT, is incapable of stimulating the transcription, despite a DNA-binding capacity (Heid et al. 2000). Additionally, mutational analysis revealed that substitution in CGTG part abolishes the binding of AhR:ARNT, thus inhibiting the transcriptional activity (Shen & Whitlock 1992).

The ligand-activated AhR binds both DNA strands at four guanine residues within the XRE recognition motif, thus interacting with the major groove of the DNA double helix (Shen & Whitlock 1992). The binding of AhR:ARNT to XRE motifs, is associated with loss of nucleosomal structure in the enhancer region and do not requires the transactivation domain (TAD). However, the protein binding to the promoter is TAD-dependent, suggesting that the transactivation domain mediates enhancer-promoter communication (Ko et al. 1997).

Similarly to some nuclear hormones described for trans-activating gene expression (*e.g.* NOR-1 (Wansa et al. 2003)) it is likely that TAD recruits co-activators and other factors that sequentially initiates the assembly of the general transcription machinery, including polymerase II complex at the promoter region (Swanson & Yang 1998; Swanson 2002). By binding XRE motifs in promoter regions, AhR regulates the expression of neighbouring genes, termed AhR gene battery (detailed in the section 1.5.3) (Fujii-Kuriyama et al. 1994).

The DNA-binding of AhR:ARNT complex, as well as the induction of some AhR-dependent genes, has been evidenced as partially mediated by phosphorylation events. The induction of AhR-dependent gene (*Cyp1a1*) following the exposure to endosulfan (a broad- spectrum cyclodiene insecticide and AhR ligand), was shown to be mediated by Ca^{2+}/calmodulin-dependent protein kinase (CaMK) and protein kinase C (PKC) in mouse hepatoma cell line Hepa-1 (Han et al. 2015). Indeed, specific kinase inhibitors reduced endosulfan-mediated CYP1A1 protein induction in Hepa-1 cells (Han et al. 2015). This suggests an implication of PKC in the induction of AhR-dependent genes. Similarly, inhibitors of protein tyrosin kinase (PTK), attenuate the ligand-induced *Cyp1a1* gene expression in rat and human hepatoma cell lines. However, the effect of PTK inhibitors seems to be dependent on the tested AhR ligands (TCDD compared to natural ligands), suggesting a ligand-specific role of PTK in the modulation of AhR signalling (Backlund & Ingelman-Sundberg 2005) (*Figure 25*).

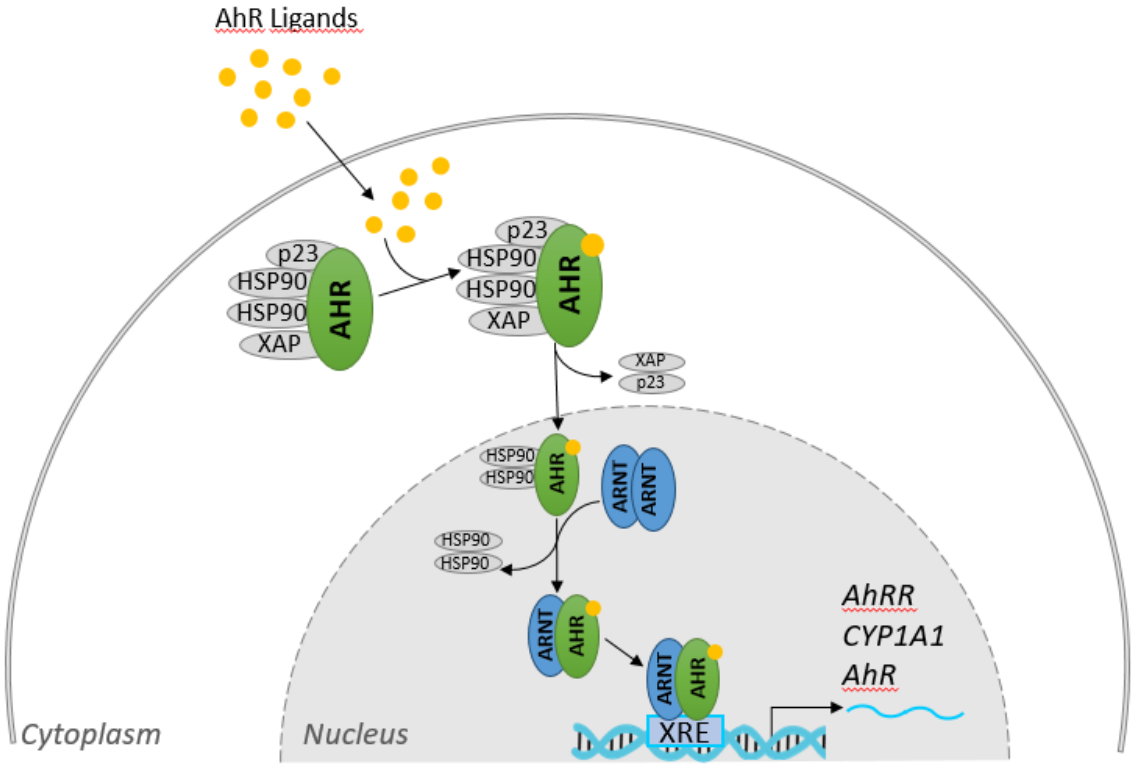


Figure 224: AhR activation signalling. *AHR: aryl hydrocarbon receptor, ARNT: AhR Nuclear Translocator, AhRR: AhR Repressor*

Ligand-induced activation of AhR signaling pathway

The AhR activation is generally mediated by the binding of diverse classes of molecules, to the ligand binding pocket in the LBD of the receptor. Although some authors reported the activation of AhR signalling in the absence of detectable exogenous ligands, the evidences for a ligand-independent activation seemed discordant as some chemicals, firstly reported to activate AhR through a ligand-independent mechanism, were lately identified as weak AhR ligands (*e.g.* omeprazole (Gerbal-Chaloin et al. 2006)). In this regards and in light of the aim of this manuscript, the herein discussion will focus only on the AhR signalling induced by ligand binding.

A huge number of molecules originated from the environment, diet as well as endogenous compounds, have been described as ligand of AhR (Barouki et al. 2007; Murray & Perdew 2017), although the physiological role in some cases remains partly unclear. The best characterised high affinity ligands for AhR include a variety of environmental contaminants such as halogenated aromatic hydrocarbons (HAHs), among which dioxin, and polycyclic aromatic hydrocarbons (PAHs). However, other molecules have been recently identified as AhR ligand and activators of AhR signalling. Consequently, AhR is not a mere dioxin-receptor but it is now recognised as more promiscuous receptor.

The activation of AhR depends upon two main characteristics of the ligand: the **affinity**, which is the property of attraction between the ligand and the receptor, and the **intrinsic effect**, referring to the receptor occupancy. A potent agonist is thus defined as a compound with both strong affinity and high efficacy. Conversely, a molecule could bind to the receptor with low efficacy, resulting in no activation of the receptor, while having a good affinity for it. This is the case of antagonist molecules, which have good or high affinity for the receptor and so compete with the agonist for the binding.

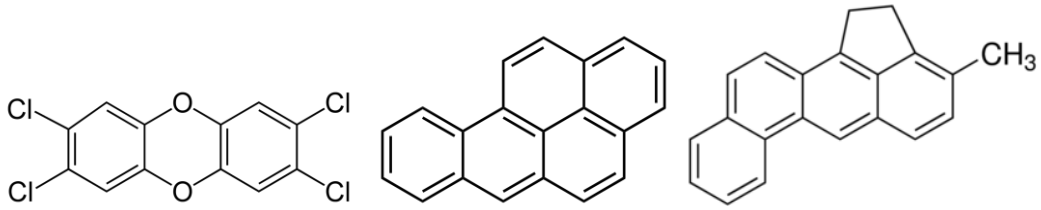
Another important characteristic of ligands is the **potency**, defined as the concentration of the ligand required to produce a detectable effect, crucial to describe the inducible effects of the specific compound and, in our contest, fundamental for evaluating the physiological impact. The variety of ligand activating AhR include not only strong agonists and antagonist but also partial agonists, which show both agonistic and antagonistic properties.

Considering that structurally diverse molecules have affinity for AhR, the intrinsic properties of the molecule seems not to offer an exhaustive classification for biological and physiological studies. Thus, in this work I would rather consider a classification based on the nature of the compounds, although other classifications exists. Considering the origin of the molecules, AhR ligands could be classified as **synthetic**, originated from anthropogenic activities and **natural** compounds, derived from biological systems and natural processed (Denison & Nagy 2003). Apart from the intrinsic properties of the molecules a differential AhR activation have been observed comparing various classes of ligands (*e.g.* synthetic and natural compounds) and the effects in distinct tissues, organs or in diverse developmental stages, suggesting ligand-specific differences in the conformation of AhR complex (Soshilov & Denison 2014). Additionally, studies on the ligand-specificity showed that AhR agonists may bind on distinct sites within the ligand-binding domain (LBD) (Petkov et al. 2010). Through a categorical COmmon REactivity PAttern (COREPA)-based structure–activity relationship model, Petkov suggested two different binding mechanisms for the ligands to accommodate in the LBD of AhR, called dioxin- and biphenyl-like, respectively (Petkov et al. 2010). The dioxin-like model predicts a mechanism that requires a favourable interaction with a receptor nucleophilic site in the central part of the ligand and with electrophilic sites at both sides of the principal molecular axis, whereas the biphenyl-like model predicted a stacking-type interaction with the aryl hydrocarbon receptor allowing electron charge transfer from the receptor to the ligand (Petkov et al. 2010). Consequently, these models could provide a possible explanation for specificity in AhR activation among different classes of ligands.

Synthetic compounds

The synthetic compounds are generally formed as by-products in the manufacturing of pesticides but also produced from inefficient combustion of organic carbon. The best characterised high affinity ligands for AhR include a variety of environmental contaminants such as polycyclic aromatic hydrocarbons (PAHs) and halogenated aromatic hydrocarbons (HAHs), among which the ,3,7,8-tetrachlotrodibenzo-p-dioxin (TCDD) is generally defined as the most powerful. These classes of hydrophilic chemicals shared a common planar structure composed by aromatic rings.

Since most HAHs are chlorinated in the lateral position, they are generally resistant to biological or chemical degradation and thus easily accumulated in food chain. HAHs like dioxins, dibenzofurans and biphenyls have relatively high binding affinity for AhR (pM to nM range); on the other hand, PAHs, as benzo[a]pyrene (B[a]P) and 3-methylcholanthrene (3MC), are characterized by a lower AhR affinity (nM to μM range) (Safe 1990; Poland & Knutson 1982). Conversely to HAHs, PAHs are compounds more easily metabolized by drug-metabolizing enzymes (CYP450 superfamily, detailed in a specific section) and thus more rapidly excreted from the body (Safe 1990; Poland & Knutson 1982).



TCDD B[a]P 3MC

Figure 235: Chemical structure of three synthetic AhR ligands.

Natural compounds

A plethora of naturally-occurring compounds are described as AhR ligands, among which the existence of endogenous AhR ligands has been suggested by numerous studies. The evidences in literature for an AhR activation in absence of exogenous molecules shed the light on the existence of endogenous ligands, although the identification of such compounds and their role at physiological concentrations remain elusive. Identified endogenous compounds include equilenin, an equine estrogen, arachidonic acid metabolites, as lipoxin 4A and heme metabolites as biliverdin, bilirubin, or hemin (Jinno et al. 2006; Schaldach et al. 1999; Sinal & Bend 1997) but more importantly, a huge class of endogenous ligands is represented by tryptophan derivatives, which will made the main object of this section (*Figure 26*).

Comparing the structure of synthetic ligands with the aromaticity of **tryptophan**, it was proposed that this amino acid or its metabolites may be endogenous agonists of the AHR. Tryptophan (Trp) is a substrate for the photolysis by UV light, originating two photoproducts displaying high AhR affinity, 6-formylindolo[3,2-b]carbazole (**FICZ**) and 6,12-diformylindolo[3,2-b]carbazole (**dFICZ**) (Rannug et al. 1995) (*Figure 26*). Competitive binding assays demonstrated that both FICZ and dFICZ have affinity for AHR comparable to TCDD, but contrarily to HAHs and similarly to PAHs, both photoproducts are substrate for drug-metabolizing enzymes (CYP450 superfamily) and thus rapidly eliminated (Rannug et al. 1987; Bergander et al. 2004; Wei et al. 1998). Additionally, a compelling body of data provides indirect evidence that the synthesis of FICZ could occur *in vivo* (Di Meglio et al. 2014; Katiyar et al. 2000).

Another Trp derivative, originated from kynurenine pathway, is the **cinnabarinic acid** shown to bind to AHR, competing with TCDD for the occupancy of the human ligand binding pocket and inducing the expression of an AhR-dependent gene (*Cyp1a1*) in zebrafish embryos *in vivo* and in human and mouse lymphocytes *in vitro* (Yamaori et al. 2015; Lowe et al. 2014).

Apart of FICZ and cinnabarinic acid, a large class of Trp derivative includes **indoles** and indole-derivatives, endogenously produced but also acquired through the diet and the metabolism of gut microbiota (refer to section 1.4.4). Among the indoles-derivatives endogenously produced, **indigo** and **indirubin**, isolated from human urine and bovine serum, shown the ability to compete for receptor occupancy with TCDD and upregulate CYP1A1 monooxygenase activity (Adachi et al. 2001) (*Figure 26*). Despite their argued role at

physiological concentrations, the enzymatic conversion of indole to the indigo precursor, 3-hydroxyindole has been demonstrated (Gillam et al. 1999). The low concentrations measured in urine seem not physiological relevant for the AhR activation, however the local concentration in different body compartment may be higher, thus allowing these compounds to reach physiologically active levels *in vivo*. Moreover, the indole-containing molecule 2-(1'H-indole-3'-carbonyl)-thiazole-4-carboxylic acid methyl ester (**ITE**), isolated from porcine lung tissue, has also been identified as AhR ligand able to compete with TCDD for binding human, murine, killifish and zebrafish AhR, to induce Cyp1a1 expression and have an *in vivo* bioactivity (Song et al. 2002; Henry et al. 2006) (*Figure 26*).

Beside the endogenously produced compounds, a massive source of AhR ligand is the **diet**. Considering the variety of dietary compounds, they probably constitute the largest and most diverse class of potential natural ligand to which humans, and animals in generals, are exposed. However, those compounds are generally rapidly metabolised and eliminated, contrarily to HAHs. Through the diet we directly introduce AhR ligands like **flavonoids**, **carotenoids**, **phenolics** and **curcumine**, derived from the consumption of tea, fruit and vegetables, that can either activate or inhibit AhR (Nishiumi et al. 2007; Zhang et al. 2003) (*Figure 26*). However, it has been demonstrated that precursors of AhR ligand or weak AhR ligands, derived from the diet, are converted in more potent AhR ligands in the mammalian GI tract (Bjeldanes et al. 1991). Indeed, the indole-3 carbinol (I3C) and 3,30-diindolylmethane, abundant in crucifers, is converted by the acidity of the stomach in indolo[3,2-b]carbazole (**ICZ**), which display one of the highest AhR binding affinities (around 0.2-3.6nM) of natural occurring products (Hooper 2011) (*Figure 26*).

Another considerable source of AhR ligands is provided by the **bacterial metabolism** (refer to section 1.4.4). Several recent studies have reported that the gut microbiota produces metabolites that potentially modulate AhR-dependent signalling (Zelante et al. 2013; Fukumoto et al. 2014; Bansal et al. 2010; Jin et al. 2014), defining another class of AhR ligands of microbial origin, produced in GI tract and not yet completely characterised. The Trp-derived **indole** is a major extracellular metabolite produced by a variety of Gram-negative and Gram-positive bacteria (refer to the section 1.4.4). Targeted metabolomic analysis of *Lactobacillus reuteri* and *Lactobacillus johnsonii* Trp catabolism has identified the indole-derivative **indole-3-aldehyde** (I3Ald) as an activator of AhR signalling *in vivo* (Zelante et al.

2013). Additionally, independent groups have identified **tryptamine** (TA) and **indole acetic acid** (IAA) as weak AhR ligands, able to stimulating AhR:XRE binding *in vivo* (Heath-Pagliuso et al. 1998; Vikström Bergander et al. 2012). However, given their weak potency as inducers of CYP1A1 expression in cell culture, it seems unlikely that IAA and TA can affect AHR signalling under physiological conditions. Nevertheless, under certain pathological states, tryptophan metabolites are likely to be relevant for the AHR signalling. Additionally, the microbial-derived **indole-3-acetate**, **3-indoxyl sulfate** (Heath-Pagliuso et al. 1998; Miller 1997; Schroeder et al. 2010; Vikström Bergander et al. 2012) and **skatole** (3-methyl indole) have been identified as AhR ligands and, skatole in particular as partial agonist (Rasmussen et al. 2016) (*Figure 26*).

Compared to the variety of commensal bacteria as well as microbial metabolites, the identified AhR-activators of microbial origins seemed a small percentage of the actual potential residing in this vast microbial community. In line with this observation, the ability of probiotic *Propionibacterium freudenreichii* in activating AhR signalling in both *in vitro* and *in vivo* studies, was described as mediated by the production of the recently identified 1.4-Dihydroxy-2-naphthoic acid (DHNA) (Fukumoto et al. 2014). Additionally, recent studies identified *Streptococcus mitis* (a commensal oral strain) and *Lactobacillus bulgaricus* OLL1181 (a probiotic strain) as activator of AhR signalling through a still unknown molecules (Takamura et al. 2011), further confirming that the identity of most of the bacteria and microbial-metabolites involved in the production of AHR-ligands are still poorly unrevealed (*Figure 26*).

Despite the general interest in human-microbiota cross-talk, some pathogens have been identified as modulators of AhR signalling pathway. In this context, *Pseudomonas aeruginosa* and *Mycobacterium tuberculosis* have been recently identified to activate AhR through the ligand binding of bacterial pigments (phenazines and naphtoquinone phthiocol, respectively) with comparable features with TCDD (Moura-Alves et al. 2014). Additionally fungi of the genus *Malassezia*, responsible for the mycoses of the skin, hair and nose, have been suggested to produce metabolites, not yet identified, able to activate AhR pathway (Magiatis et al. 2013; Mexia et al. 2015) (*Figure 26*).

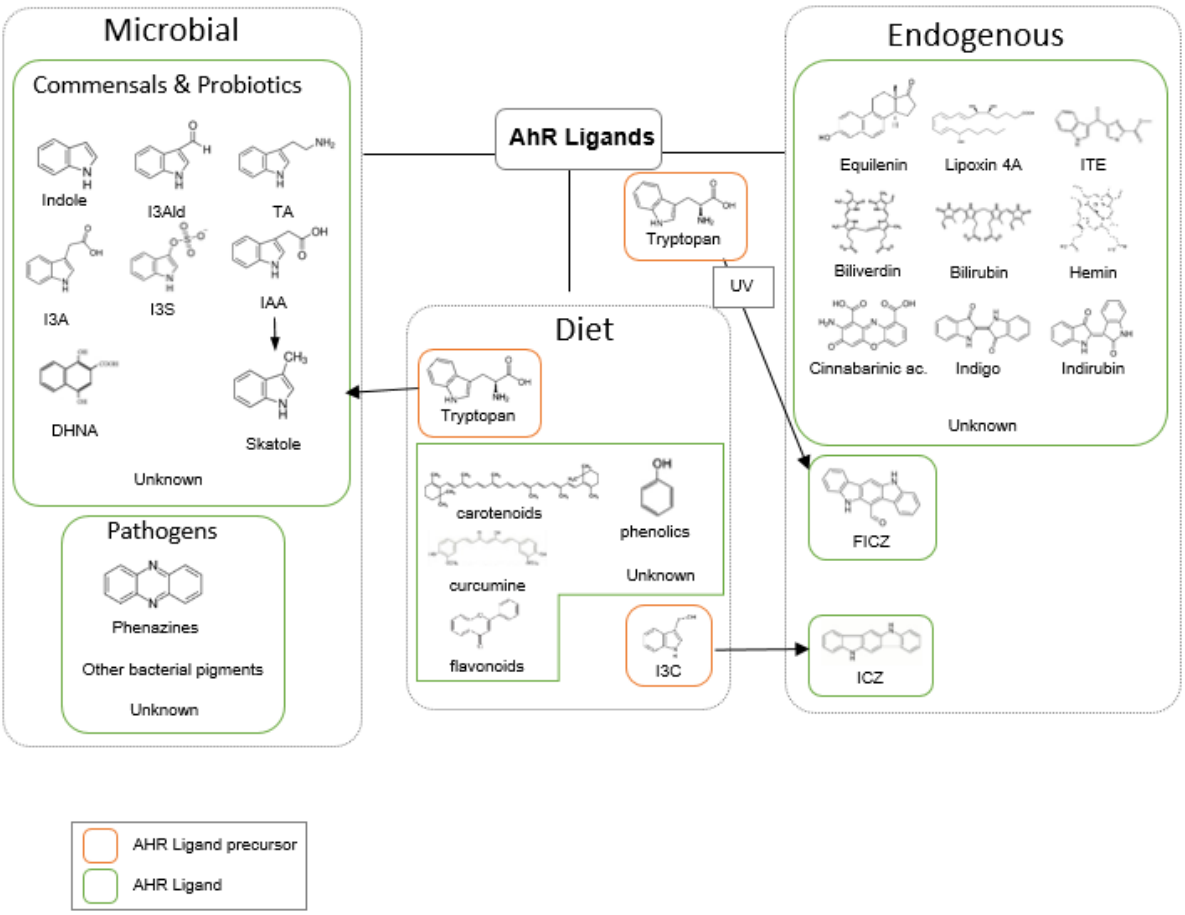


Figure 246: Structures and origins of some natural AhR ligands.

The abundance and variety of AhR activators evidenced a broad sensing potential of AhR in responding to exogenous and endogenous compounds that, by binding to the LBD, stimulate the expression of AhR-dependent genes that impact host physiology.

AhR-Dependent Genes and Negative Feed-Back Regulation of the AhR Signalling

The cytochrome P450 1A1 (CYP1A1), a drug metabolizing enzyme, was the first described as dependent on AhR activation signalling. However, later on, AhR was found to regulate the expression of a variety of other drug metabolizing enzymes, including cytochrome P450 1A2 (CYP1A2), cytochrome P450 1B1 (CYP1B1), glutathione-S transferase (GST) A1, NAD(P)H:quinone oxidoreductase 1 (NQO1), UDP-glucuronosyltransferase (UGT) 1A, aldehyde dehydrogenase(ALDH)3A1. Nonetheless, the list of genes influenced by AhR activation continues to grow. Indeed, recent works suggested the impact of AhR on PPARγ- and estrogen receptor (ER)- related gene expression, on the gene-encoding aromatase cytochrome P450 19A1 (CYP19A1) and on enhanced expression of genes encoding epidermal growth factor(EGF)-like proteins (Haarmann-Stemmann et al. 2009; Kerkvliet 2009; Puga et al. 2009; Beischlag & Perdew 2005).

Interestingly, AhR also regulates the expression of its repressor AhRR to ensure a negative regulator loop. Indeed multiple mechanisms evolved to suppress sustained AhR activity, implying that the prolonged receptor signalling is physiologically deleterious. Among these regulatory mechanisms, the AhR-dependent expression of AhRR as well as CYP1A1 constitute a negative regulatory feedback loop through the metabolic depletion of exogenous and endogenous ligands.

Cytochrome P450 Expression and Regulation of AhR Signalling

Cytochrome P450 (CYP450) is a superfamily of enzymes characterized by a single polypetide chain ranging in size 40-50kDa and containing a single heme group coordinated to a cysteine residue (*Figure 27*).

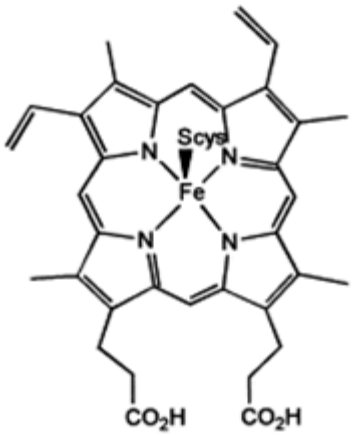


Figure 257: Active site of Cytochrome P450

Since the first identification in the 80s, hundreds of different P450 genes have been identified making this enzyme family one among the most widespread and diverse biocatalysts.

CYP enzymes are widely distributed in different organisms including animals, fungi, bacterial and virus (Seliskar & Rozman 2007). In mammals, up to 18 CYP450 families, have been identify to encode for 57 genes in the human genome (Nebert et al. 2013). However, even within mammals, not all the organisms have the same set of P450 enzymes especially the ones involved in the metabolism of some exogenous compounds, reflecting the evolutionary adaptation to chemicals exposure of different species (Nelson et al. 2004).

CYP450s catalyze the monooxygenase reaction of various exogenous (xenobiotics) and endogenous substrates. In human, CYPs are typically anchored to the endoplasmic reticulum by an *N*-terminal domain within the catalytic domain residing in the cytosol, partially embedded in the membrane. To accomplish its function in metabolizing endogenous and exogenous compounds, CYP needs a source of electrons, provided by its redox partners: the cytochrome P450-reductase (CPR), a large flavoprotein (approximately 77 kDa) containing FAD and FMN domains, and cytochrome b_5 (b_5), a small hemo-protein (15 kDa) which promotes catalysis through electron transfer and allosteric regulation (Barnaba et al. 2017).

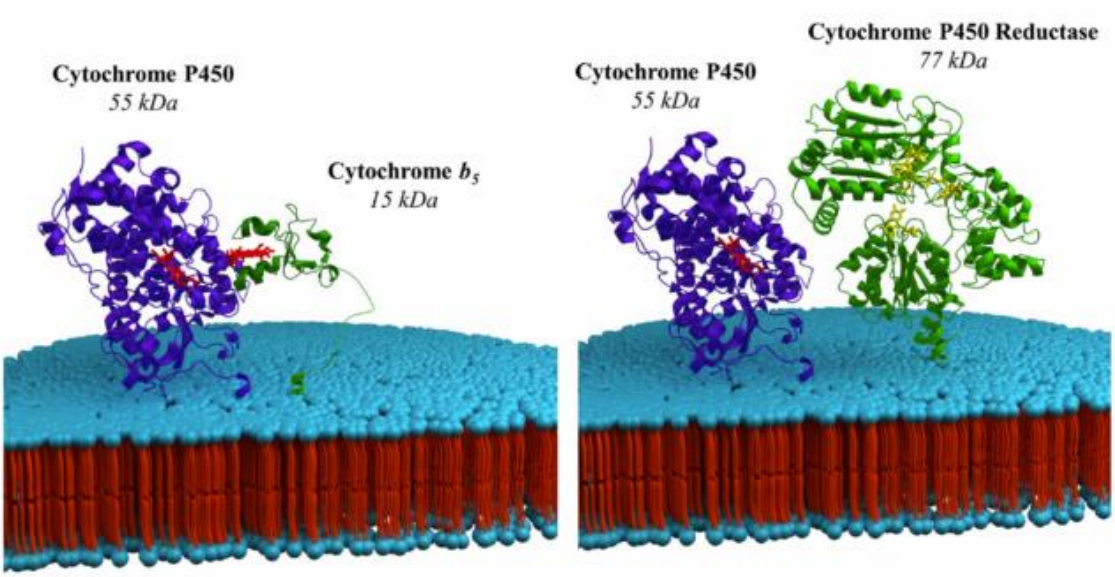


Figure 268: Model structure of cytochrome P450 with cytochrome b5 (left) and cytochrome P450 with cytochrome P450-reductase (right) in a lipid bilayer *(Barnaba et al. 2017). Models constructed from the crystallographic or solution-state nuclear magnetic resonance*

CYP450 was firstly described as 'hepatic drug detoxification system' involved in phase I metabolism of xenobiotic. The xenobiotic metabolism, divided in two phases (phase I and II), consists in chemical modifications of the xenobiotic molecule (*e.g.* hydroxylated, deaminated, oxidated or dealkylated) in phase I and a conjugation with water-soluble molecules, in phase II (*e.g.* UDP-glucuronic acid, acetyl-coA, glycin or sulphates) to inactivate and facilitate its elimination. Beside this firstly described role in detoxification, CYP450 is now known to include uncounted enzymatic reactions, among which arachidonic acid metabolism and eicosanoid biosynthesis, cholesterol, sterol and bile acid biosynthesis, steroid synthesis and catabolism, and vitamin D3 synthesis and catabolism. Moreover, later studies evidenced the ubiquitous expression of CYP450s in human tissues, other than the sole liver, including the gastrointestinal tract.

Ligand-induced activation of AhR induces the expression of cytochrome P4501 (CYP1) enzymes through the binding of AHR:ARNT complex to the XRE motif in the region of their promoter. In turn, CYP1 enzyme oxygenate AHR ligands, leading to their metabolic clearance and detoxification, thereby terminating AHR activation. The ligand clearance emerged to be essential in the AhR signalling, indeed prolonged activation either by ligands that resist metabolic clearance or by constitutively active AHR, has deleterious effects, such as carcinogenesis (Chang & Puga 1998; Mitchell & Elferink 2009). Consequently, multiple

mechanisms evolved to suppress the sustained AhR activation, among which the negative feedback assured by enzymes of the CYP1A and CYP1B subfamilies (CYP1).

Low affinity agonists are generally characterized for being metabolic labile molecules (*e.g.* PAHs, natural compounds and microbial metabolites) and substrates for cytochrome enzymes. Thus, the AhR activation induced by these labile agonists, stimulates a transient signalling that not eliciting toxic responses. However, an excessive ligand clearance induced by a constitutive active Cyp1a1 was shown to be deleterious. Indeed, constitutive expression of *Cyp1a1* throughout the body or restricted specifically to intestinal epithelial cells (IECs) produces a quasi *Ahr*-deficient state that lead to a loss of AHR-dependent ILC3 and T_H17 cells and increased susceptibility to enteric infection (Schiering et al. 2017). These evidenced suggests that IECs serve as gatekeepers for the supply of AHR ligands to the host and emphasise the importance of feedback control in modulating AHR pathway activation.

However, many high affinity AhR agonists, such as HAHs and in particular dioxins, are poor substrates for these enzymes, inducing a sustained AhR stimulation that culminates in toxic responses. Thus, to counterbalance the prolonged stimulation, an additional control mechanisms take place, involving the AhR-dependent expression of its repressor AhRR.

AhRR Expression and Regulation of AhR Signalling

AHRR was isolated and identified in 1999 from a mouse intestinal cDNA library and firstly described as an "AHR-like" protein, hypothesized to compete for binding to ARNT and work as a negative regulator of the receptor activity (Kawajri & Fujii-Kuriyama 2007; Baba et al. 2001; Mimura et al. 1999). The AhRR was then better characterized as a highly evolutionarily conserved bHLH/PAS protein in vertebrates, closely related to the AHR (Hahn et al. 2009; Mimura et al. 1999; Haarmann-Stemmann et al. 2007) and localized in chromosome 13C2 in mouse, 1p11 in rat and 5p15.3 in human gene (Baba et al. 2001). Orthologous proteins have been described in several mammalian species, including mice, humans, rats, amphibians, and many types of bony fish (Tsuchiya et al. 2003; Korkalainen et al. 2001).

In humans, as well as in mouse and rats, AhRR expression has been detected in numerous tissues and cell lines revealing a tissue- and cell-type specific expression (Tsuchiya et al. 2003; Korkalainen et al. 2001). Notably, in adult human, AHRR was detected in liver, breast, colon, kidney, lung, bladder, uterus, ovary, adrenal gland, with a prominent high expression in testis. Furthermore, AhRR mRNA expression was also evidenced in the lung, kidney, spleen, and thymus of the human fetus, but not in brain, liver, heart, and muscle (Tsuchiya et al. 2003). This observed embryo-fetal expression, suggests a possible developmental role for the AHRR, although in one study, AHRR$^{-/-}$ mice are fertile and the offspring appear to develop normally (Hahn et al. 2009; Hosoya et al. 2008).

Studies on AhRR structure revealed high sequence similarity with the AhR in the *N*-terminal region in which are residing both NLS and NES, homologous to AHR (Kawajiri & Fujii-Kuriyama 2007). However, aside the *N*-terminal region, the repressor significantly diverges from AHR towards the *C*-terminus (Kewley et al. 2004). Owing to its structural similarity to the AhR in the *N*-terminal half, the AhRR can also dimerize with ARNT and bind to XREs. However, since the *C*-terminal part of the AhRR protein lacks the transactivation domain (TAD) (Baba et al. 2001; Mimura et al. 1999), the DNA binding is proposed to recruits co-repressors and to function as a transcription repressor of AhR activation (Mimura et al. 1999). Additionally, AhRR is also lacking the PAS-B motif and LBD, making the protein unable to bind ligands (Baba et al. 2001; Mimura et al. 1999) (*Figure* 29).

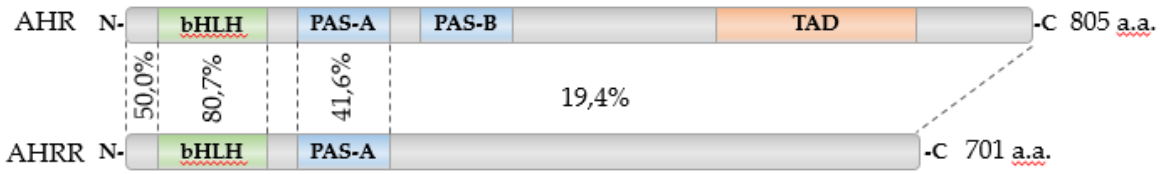


Figure 279: Domain structure of mouse AHR and AHRR. *The percentage of identity between AHR and AHRR are indicated.*

As expected from the inducible expression of AHRR in response to an AHR activation, the *Ahrr* gene has three (in mouse) or four (in human) copies of the XRE sequence in the promoter.

The AhR-dependent induction of AhRR may serves as a complementary regulatory mechanism to abolish the sustained AhR stimulation induced by ligands resistant to CYP metabolization, such as HAHs. The hypothesis for the AhRR negative feedback mechanism was supported by the observed structural similarity between AhR:ARNT and AhRR:ARNT heterodimers as well as the reported increase in AhRR mRNA expression in mice expressing a constitutively active AhR (Andersson et al. 2002; Sakurai et al. 2017). To gain insights into the transcriptional repression mechanism of AhRR, Sakurai and co-workers recently analysed the crystal structure of the AhRR:ARNT complex compared to AhR:ARNT. The AhR residues interacting with XRE DNA, are conserved in AhRR, as well as the similar positioning of bHLH domain in AhRR:ARNT compared to AhR:ARNT complex. These observations bring to the hypothesis that AhRR:ARNT efficiently compete with AhR-ARNT for DNA binding (Sakurai et al. 2017) (*Figure 31*).

However, it was observed that ARNT overexpression failed to reverse the AhR repression induced by AhRR, suggested that the inhibition does not occurred solely by sequestration of ARNT. Moreover, it was described that mutated AhRR, not able to bind XRE, was remain functional, thus implying an inhibitory role independent on XRE binding. These evidence opens the way to a second hypothesis of a "transrepression" of AhR signalling through protein-protein interactions rather than by inhibition of AHR-ARNT complex (Evans et al. 2008) (*Figure 31*). Consequently, different authors have proposed a scenario depicting AhRR activity as functionally interfering with AhR transformation, specifically on the receptor dissociation from the chaperone proteins inside the nucleus and dependent on the recruitment of yet undescribed co-repressor proteins and independently of ARNT binding (Mitchell & Elferink 2009; Evans et al. 2008).

Additionally, SUMOylation of the C-terminal region of AhRR has been described as critical for recruiting various co-repressor molecules to the promoter region, thus leading to a third hypothesis for the AhRR repression of AhR signalling, through the recruitment of SUMO E3 ligase for the SUMOylation of AhRR/ARNT, thereby down-regulating the transcription (Oshima et al. 2009; Sakurai et al. 2017) (*Figure 31*). Collectively, these data suggest that the current model of AHRR-mediated repression is not fully understood and might be composite, thus additional studies are necessary to completely decrypt this repressor mechanism.

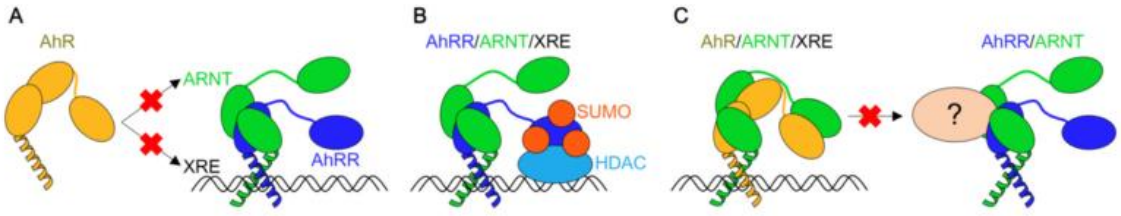


Figure 30: Models of AhRR transcriptional repression mechanism:
A. Competitive repression model. *AhRR competes with AhR for heterodimerization with ARNT and bindind to XRE DNA.* ***B. Corepressor-mediated repression model.*** *AhRR-ARNT heterodimer binds to XRE DNA and recruits coreoressors, which leads to transcription repression.* ***C. Transrepressor model.*** *AhRR:ARNT heterodimer competes with AhR:ARNT for dinding to unknown interaction proteins (Sakurai 2017)*

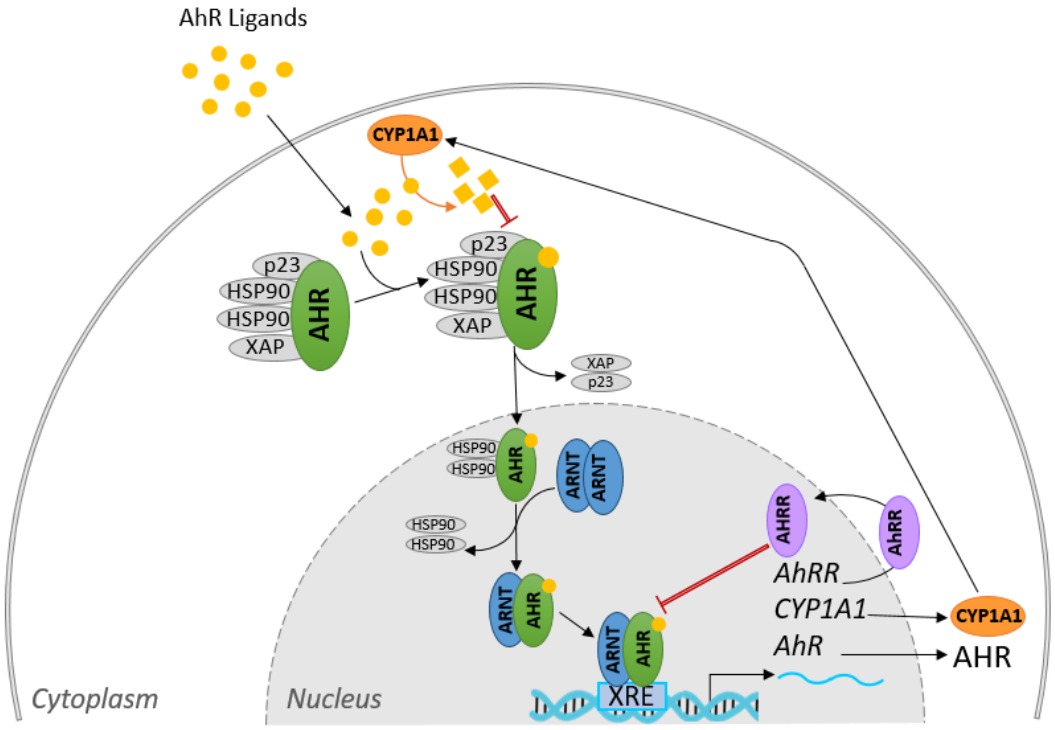


Figure 31: Schematic representation of the AhR regulation through CYP450 and AHRR.

Other Regulators of the AhR Signalling Pathway

TiPARP

Beside of AhRR, MacPherson has recently reporter that TiPARP (also known as ARTD14), a mono-ADP-ribosyltransferase, that is a ligand-induced negative regulator of AHR transactivation. Similarly to AHRR, TiPARP seems to be part of an auto-regulatory negative feedback loop regulating AHR activity (MacPherson et al. 2014) (*Figure 32*).

Comparing the ability of AHRR and TiPARP to repress AHR signalling, a number of similarities emerged. In particular expression of both AhRR and TiPARP is induced by ligand-activated AHR and the overexpression of ARNT did not impact the repression activity. However, TiPARP overexpression increases the proteolytic degradation of AHR, suggesting that it may act as a more general regulator of AHR activity by a mechanism different from AHRR one (MacPherson et al. 2014).

However further studies are necessary to dissect the overlapping and independent mechanisms through which AHRR and TiPARP repress AHR.

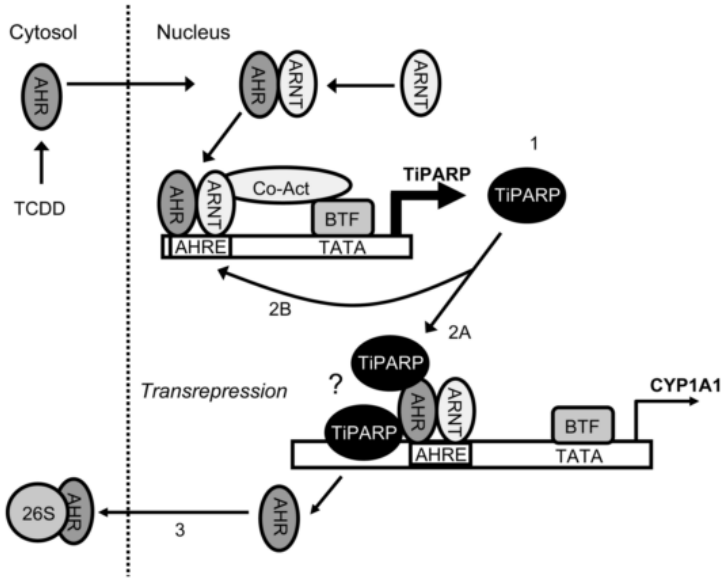


Figure 282: Proposed model of repression of AHR signalling by TiPARP. *(1). TiPARP is recruited to the AHR:ARNT bound AHRE(XRE) of targets to repress transcription (2A). TiPARP represses its own transcription by binding to AhR:ARNT (2B) AHR dissociates from AHRE (XRE motif), translocates from the nucleus and is proteolytically degraded by the 26S proteasome (3). BTF, basal transcription factors; Co-act, co-activators. (MacPherson 2013)*

Ubiquitination

Beside the aforementioned regulatory mechanisms, AhR has been shown function as a ligand-dependend E3 ligase, targeting itself and substrate proteins such as steroid receptors and β-catenin for proteosomal degradation (Ohtake et al. 2007; Ohtake et al. 2009). The AhR-associated ubiquitin ligase complex has been biochemically purified from HeLa cells (Ohtake et al. 2007). The characterized complex includes cullin 4B (CUL4B), damaged- DNA-binding protein 1 (DDB1), and Rbx1 together with subunits of the 19S regulatory particle (19S RP) of 26S proteasome as well as ARNT and transducin-beta-like 3 (TBL3) (Fig.) (Ohtake et al. 2007). The core of this ubiquitin ligase complex appears to constitute a CRL-type E3 ligase, belonging to the Cullin/RING ubiquitin ligases (CRL) subfamily of ubiquitin ligases. In line with this, the AhR-associated ubiquitin ligase complex was termed $CUL4B^{AhR}$ (*Figure 33*).

Complex assembly and ubiquitin ligase activity of $CUL4B^{AhR}$ *in vitro* and *in vivo* are dependent on the AhR ligand. In the CUL4BAhR complex, ligand-activated AhR acts as a substrate-specific adaptor component that targets substrate proteins for degradation. Coherently, AhR levels are reduced upon TCDD treatment indicating that ligand binding induce a self-ubiquitination of the receptor (Kawajiri et al. 2009; Ohtake et al. 2007), although this effect seems to be dependent on the ligand. Altogether these evidences uncovered an additional function for AhR and demonstrate a non-genomic signalling pathway in which fat-soluble ligands regulate target-protein-selective degradation through an ubiquitin ligase complex. Additionally, through the $CUL4B^{AhR}$ complex, AhR seems able to modulate the expression, function and activity of other proteins in response to exogenous and endogenous stimuli.

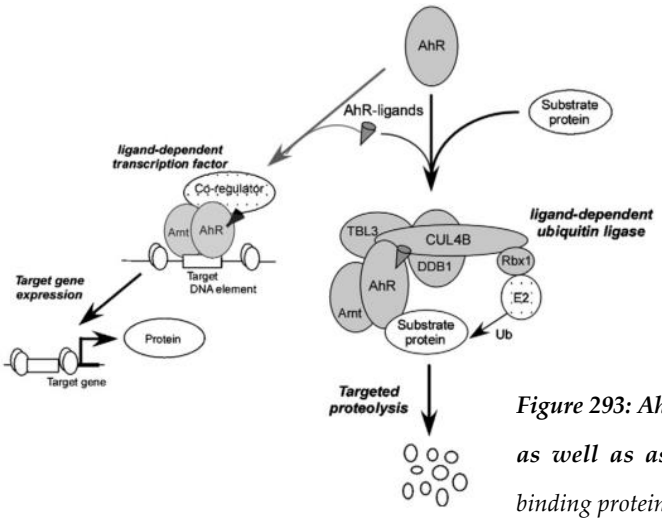


Figure 293: AhR serves as a ligand-dependent ubiquitin ligase, as well as as a transcription factor. *DDB1, damaged-DNA-binding protein 1; TBL3, transducin-beta-like 3. (Ohtake 2007)*

In conclusion, the nature of the AhR ligand and, in particular, its metabolic stability determines its sensibility or resistance to metabolic clearance thus inducing transient or sustained signalling, respectively. Consequently, transient and sustained AhR activation induce divergent physiological effects with mainly toxic consequence in the sustained stimulation. In this context, a recent study conducted by Tijet *et al.* demonstrated that over 200 distinct genes were differentially expressed in the livers of $AhR^{-/-}$ and wild type mice at steady state (Tijet et al. 2006). This suggests that transient signalling induced by endogenous ligands, contributes to numerous physiological responses, among which reproduction, growth and development (Tijet et al. 2006). The up-regulation of some genes, induced by endogenous ligands, was not reproduced by TCDD treatment indicating that sustained receptor activation does not faithfully replicate the response induced by the transient signalling. In line with this it is conceivable that less responsive XRE-regulated genes could have increasingly probability to recruit functional receptor complex during prolonged stimulation of the receptor rather than a transient activation. Consequently, this could lead to a transcriptome profile different between prolonged and transient signalling (Mitchell & Elferink 2009). Hence, the duration of receptor signalling influences both quantitative (*e.g.* amount of gene expression) and qualitative (*e.g.* gene expression spectrum) parameters, with a consequent disequilibrium from homeostatic toward a toxic response and important consequence at physiological level.

Beside the duration of the AhR stimulation, important species-specific differences have been evidenced regarding the ligand binding, which could strongly impact the physiological effect of the receptor activation. Studies conducted on mAhR and hAHR affinity for TCDD, evidenced a 10-fold decreased affinity for TCDD in human compared to mice (Flaveny et al. 2009). Considering the intra- and inter-species differences in AhR homologs within the animal kingdom and, in particular between human AhR (hAhR) and mice AhR (mAhR) (refer to the animal section), it is seems conceivable that protein homologs could have different binding affinities although detailed differences of the molecular arrangement within the LBD were not available until recently. By molecular docking and using the prototypical AhR ligand TCDD, Bisson and co-workers evidenced differences in the binding energy, even if the docking of TCDD in mAhR and hAhR is similar in the binding pocket, thus confirming the higher affinity of mAHR for TCDD compared to hAHR, likely due to the mouse Ala375 residue replaced by a Val in hAhR (Bisson et al. 2009)(*Figure 34*).

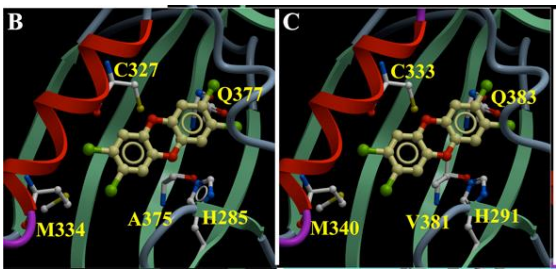


Figure 304: TCDD docking orientation into mouse (B) and human (C) AhR-LBD binding pocket. *The protein backbone is displayed as ribbon and colored by secondary structure. The residues are displayed as sticks and colored by atom type with the carbon atoms in white. TCDD is displayed as sticks and colored by atom type (carbon atoms in yellow) (Bisson 2009).*

Additionally, a recent *in silico* modelling analysis, predicts the structure-activity selectivity of indole associated with hAHR compared to mAHR (Hubbard, Murray, Bisson, et al. 2015). Based on the established specificity of hAHR for indirubin, which closely resembles two covalently linked indole moieties, Hubbard and co-workers described that hAHR ligand binding pocket can adopt an energetically favourable conformation that is permissive for two molecules of indole, whereas the mAHR is more restrictive, allowing only a single indole to bind, resulting in very weak agonist activity and thus explaining the higher affinity of hAHR for indoles (Hubbard, Murray, Bisson, et al. 2015)(*Figure 35*).

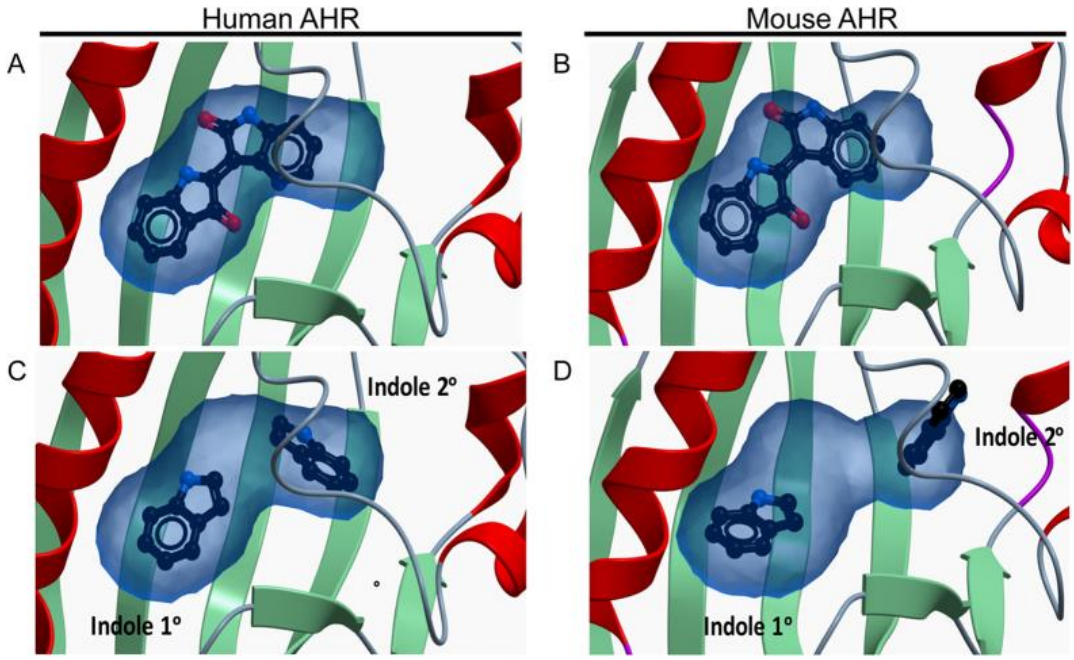


Figure 315: In silico modelling of AHR ligand binding domain. *Homology modelling of indirubin optimized ligand binding in (**A**) hAHR and (**B**) mAHR; The predicted two indole-binding model in (**C**) hAHR and (**D**) mAHR ligand binding domain. Blue shading indicates the space-filling volume of the ligand binding pocket (Hubbard 2015).*

In light of these inter-species differences as well as the plethora of ligand classes and their abundance in the intestine, it appears essential to critically evaluate the results obtained on animal models, considering the limitations in the extrapolating a physiological role in human, at least concerning AhR pathway. However, the development of a transgenic mice expressing hAHR (Flaveny et al. 2009) constitutes an extremely useful tool to eliminate several of these challenges. Coherently with the results on the receptor affinities, these hAHR mice showed a differential TCDD-mediated toxicity compared to the mAHR counterpart. Additionally, hAHR transgenic mice evidenced the higher affinity of human AhR for certain compounds, including indirubin and quercetin, compared to mAHR, confirming that AHR may be distinctly regulated in a species-specific fashion (Flaveny et al. 2009). Consequently, hAHR mice emerged as an extremely promising tool for the study of human AhR in animal model, avoiding the inter-species differential activations. This model could allow the study of human-specific AhR ligands, their physiological roles as well as their interaction with other metabolic pathway *in vivo*. Additionally, a mouse model expressing the human AhR could be a better transposable model for the influence of gut microbiota in human physiology through the AhR signalling. Indeed, considering the differences described in ligand affinity between human and mouse AhR, it is conceivable that some microbial-derived metabolites could display a differential activation in mouse compared to human and thus a different physiological impact.

1.5.4. Physiological role of AhR

Identified as xenobiotic receptor, AhR is part of the detoxification system of the organism. Vertebrate AhR is the only able to bind structurally diverse compounds and sense environmental stress, contrarily to invertebrate AhR orthologs (Butler et al. 2001). At the time of its discovery, AhR made the object of intensive toxicological studies that brought to the description of its role in the up-regulation of phase I and II biotransformation enzymes, following the binding of environmental contaminants. Additionally, the effects induced by the exposure to dioxins have been described as mediated by the binding of these chemicals to the AhR. The dioxin-binding results in a prolonged AhR stimulation that culminates in multiple toxic effects including thymus atrophy, liver enlargement, wasting, cardiovascular diseases, chloracne, tumorgenesis and immune suppression (Andersson et al. 2002). Coherently, mice with a constitutive active AhR, showed similar signs of toxicity and $AhR^{-/-}$ mice confirm that the effect is clearly mediated by AhR signalling (Andersson et al. 2002; Mimura et al. 2003). Consequently, the discrepancy between invertebrate and vertebrate AhR, was proposed as an evolutionary signature in higher organism for sensing anthropic environmental pollutants and detoxification from other xenobiotic molecules. However, in the last decades, several publication have described the downstream effect of AhR signalling in diverse aspects of mammalian biology that extend beyond (or are entirely independent from) the xenobiotic metabolism.

The expression of the AhR during early embryonic development, the ubiquitous expression in adult organisms and the strong conservation throughout evolution, suggest a role of the receptor in physiological process as well as in pathological conditions (Abbott et al. 1995; Peters & Wiley 1995; Esser & Rannug 2015; Hahn 2002).

Studies on the invertebrate AhR orthologs ahr-1 (*Caenorhabditis elegans*) and spineless (*Drosophila melanogaster*) revealed functions in the control of neuron differentiation in the nematode and neuronal development in fly, especially in dendrites, antenna, tarsus and ommatidia development (Huang et al. 2004; Wernet et al. 2006). Similarly, AhR deficient mice display vascular abnormalities in hearth, uterus and in the liver. In particular, in liver reduced size, patent ductus venous, portal fibrosis and steatosis have been observed (Fernandez-Salguero et al. 1997). In uterus, vascular mineralization with thrombosis was described. Coincident with difficulties in maintaining pregnancy, the uterus abnormalities suggested that

AhR signalling has an additional role in the homeostasis of female reproductive system (Fernandez-Salguero et al. 1997). AhR deficiency have been reported to impact the development and maintenance of the immune system. A quantitative comparison of splenic lymphocytes number between $AhR^{-/-}$ and $AhR^{+/+}$ mice indicates a reduced amount of lymphocytes when AhR is depleted, although this difference appears to be age-dependent. In particular, $AhR^{-/-}$ mice at 2 to 3 weeks of age contained 75 to 85% fewer lymphocytes than their littermates $AhR^{+/+}$ (Fernandez-Salguero et al. 1995). As $AhR^{-/-}$ mice aged, the number of lymphocytes increased up to approximately normal number by 10-12 weeks after birth, and then it further decrease along with ageing (-50% less lymphocytes at 25-32 weeks compared to $AhR^{+/+}$)(Fernandez-Salguero et al. 1995). Furthermore, $AhR^{-/-}$ mice showed a reduced number of lymphocytes in peripheral lymph node, although the ratio between T and B lymphocytes remained normal (Fernandez-Salguero et al. 1995). Consequently, a role of AhR in the maintenance and development of immune system emerged. In the gut, $AhR^{-/-}$ mice bore caecal lesions with moderate or high grade of atypia, adenoma, and adenocarcinoma with a consequent implication in gastrointestinal physiology (Kawajiri et al. 2009). Similarly, epithelial lesions have been observed in skin (Fernandez-Salguero et al. 1997), indicating a role of AhR in epithelial maintenance. Coherently, at cellular level, depending on the context, AhR signalling influences several host responses and pathways including cell cycle (Ma & Whitlock 1996), the regulation of cellular differentiation for keratinocytes, hematopoetic cells, adipocytes and immune cells (Haas et al. 2016; Shimba et al. 2003; B. W. Smith et al. 2013) as well as the response to antioxidants and hormone-like estrogen (Lee et al. 2003).

Altogether these evidences show the role of AhR in a myriad of cellular and physiological processes, even if some effects have to be assumes as ligand- and/or species-specific (discussed in the previous section).

AhR in Intestinal Homeostasis

AhR expression differs significantly between tissues but it is conspicuous in barrier organs such as the gut in which AhR is expressed by intestinal epithelial cells (IECs) and by different immune cells (Esser & Rannug 2015; Chmill et al. 2010). Several reports evidenced that ligand-induced AhR activation is involved in the development, function and maintenance of the physical and immunological intestinal barriers, although in some cases the exact mechanism still remains elusive.

Role of AhR on intestinal epithelial barrier function

The proliferation and maintenance of the intestinal epithelium is the front line physical barrier for the protection toward environmental contaminants as well as potential harmful microbes. $AhR^{-/-}$ mice showed an impaired proliferation of colonic crypt stem cells with a consequent defect in cell renewal and integrity of the epithelium (Stockinger et al. 2014). Additionally, an enhanced apoptosis of epithelial cells was observed when AhR was depleted in mice (Chinen et al. 2015). The role of AhR on epithelial integrity was further described to be mediated by the ligand-dependent activation of the receptor. Indeed, AhR agonist FICZ in *in vivo* models ameliorates intestinal obstruction- and DSS- driven intestinal permeability by rescuing the expression of epithelial tight junction (TJ) proteins zonula occludens-1 (ZO-1), Occludin and Claudin-1 (Yu et al. 2018; Han et al. 2016). This ligand-dependent effect of AHR on epithelial permeability was confirmed by *in vitro* experiments in which Caco-2 cells evidenced significant change in TJ distribution following to FICZ exposure (Yu et al. 2018). In this context, a recent study showed that loss of Notch1 signalling in Caco-2 cells, counteracted the development of TJs induced by FICZ (Liu et al. 2018); consequently, a mechanism mediated by the up-regulation of Notch1 was proposed for the AhR-mediated protection against intestinal damages. Similarly, bacterial derived indoles has been shown to increases epithelial TJ integrity and reduce intestinal inflammation (Bansal et al. 2010), although it is still not clear if this effect is mediated by AhR.

Despite its significance at epithelial level, it should be point out that AhR is expressed in a variety of other cells, thus the effects on the epithelium could be mediated by signalling through other cell types. For example, the AhR-dependent production of IL-22 by immune

cells (ILCs and TH17), have a major effect on IEC proliferation and production of antimicrobial peptides (Kiss et al. 2011; Qiu et al. 2012). In turn, IEC may modulate the effect induced by AhR on other cell types. Indeed, constitutive CYP1A1 activity in IECs, but not in adaptive immune cells, restricts the availability of AHR ligands to cells in the intestinal lamina propria, resulting in loss of AHR ligand-dependent ILC3 cells (Schiering et al. 2017). In turn, the reduction in ILC3 affect the level of IL-22 in the colon with a consequent increased sensitivity toward *Citrobacter rodentium* infections (Schiering et al. 2017). Conversely, constitutive expression restricted to adaptive immune cells showed normal number of ILC3 in steady state and a survival rate to *Citrobacter rodentium* infection similar to wild type mice (Schiering et al. 2017). Consequently, this suggest a major role of IECs in molecular clearance, fundamental for the availability of AhR ligands to other cell types and a consequent control of the AhR signalling (Schiering et al. 2017).

Aside from the epithelial proliferation, the barrier function in intestine is also guaranteed by the production of mucus. In a lung epithelial cell line, AhR activation by TCDD was shown to induced the expression of MUC5AC (Wong et al. 2010). Considering that MUC5AC is also expressed in human intestine (Guyonnet Duperat et al. 1995), it is conceivable that similarly to what observed in the lung, AhR could be involved in the modulation of mucus production in the intestine. However, to my knowledge not such studies have been reported.

Role of AhR on intestinal immune response

A second line of protection toward potential environmental threats, is guaranteed by the intestinal immune cells. AhR is expressed by different immune cells, including intraepithelial lymphocytes (IELs), T_H17 cells ,Treg, innate lymphoid cells (ILCs), macrophages, DCs and neutrophils (Li et al. 2011; Qiu et al. 2013; Kimura et al. 2009; Nguyen et al. 2010; Zindl et al. 2013; Frericks et al. 2007). Consequently, AhR signalling emerged to be pivotal also in the regulation of the intestinal immune homeostasis.

Under the steady state, tissue resident Treg in the gut express high levels of AhR (Ye et al. 2017) that may represent a mechanism of tissue adaptation, rendering gut Tregs readily activated by environmental cues (*e.g.* ligands) to exert their suppressive function locally for gut homeostasis. Ligand-activated AhR was described to have a major role in the thymocyte linear decision as well as in shift in immune-cell subset. In particular, by modulating Stat-1-mediated cytokine expression, AhR influences the functional differentiation of naïve T-cells into regulatory T cells (Treg) and T-helper 17 (T_H17) cells, involved in autoimmunity and defence against microbial infections (Veldhoen et al. 2009; Veldhoen et al. 2008; Quintana et al. 2008; Kimura et al. 2008). In T_H17 cells, like in ILC3s, AhR induces the production of the critical interleukin-22 cytokine involved, among other functions, in the epithelial proliferation and AMP production to enhance the intestinal epithelial barrier (Qui et al. 2013). Both endogenous and exogenous ligands of the AhR, such as FICZ and dioxin, have been described to influence its activity; however, different types of ligands could generate divergent effects, reinforcing the evidences for a ligand-specific receptor plasticity.

Additionally, AhR has been evidenced to act as a molecular switcher of monocyte differentiation toward monocyte-derive-DCs (Goudot et al. 2017), which are major producers of IL- 23 and inducers of T_H17 cells (Segura et al. 2013).

AhR is also a critical modulator in maintaining the number of intraepithelial lymphocytes (IELs) in the intestine where a receptor deficiency or the depletion of AhR ligands compromises the maintenance of IELs. Indeed, the number of small intestinal TCRγδ+ IELs was reported to increase when mice are fed with a diet supplemented in the AhR ligand I3C (Li et al. 2011), suggesting an important link between AhR, diet and modulation of immune populations in the gut. Additionally, the capacity of IELs to produce antimicrobial peptides (AMPs) following to bacterial exposure (Ismail et al. 2011) is decreased when AhR

activity is reduced in mice, resulting in an increased bacterial load, particularly associated with an enhanced contribution of species of the *Bacteroidetes* phylum (Li et al. 2011).

Moreover, the AhR expression was also detected in B cells in which endogenous AhR ligands were described to have a role in maintaining the functional response of B cells to antigen activation (Villa et al. 2017). AhR deficiency in B cell results in a reduced proliferation and in a lost out in competition with AhR-sufficient B cells, both in settings of homeostatic proliferation and upon antigen-driven proliferation (Villa et al. 2017).

The expression of AhR in a huge variety of cell types evidences its huge potential in influencing the most diverse cellular responses with consequences for the entire organism that starts to be described at physiological level. Although the AhR activation appeared to be fundamental at basal homeostatic state, its protective role was largely evidenced in pathological conditions, among which intestinal inflammation and pathogen infections.

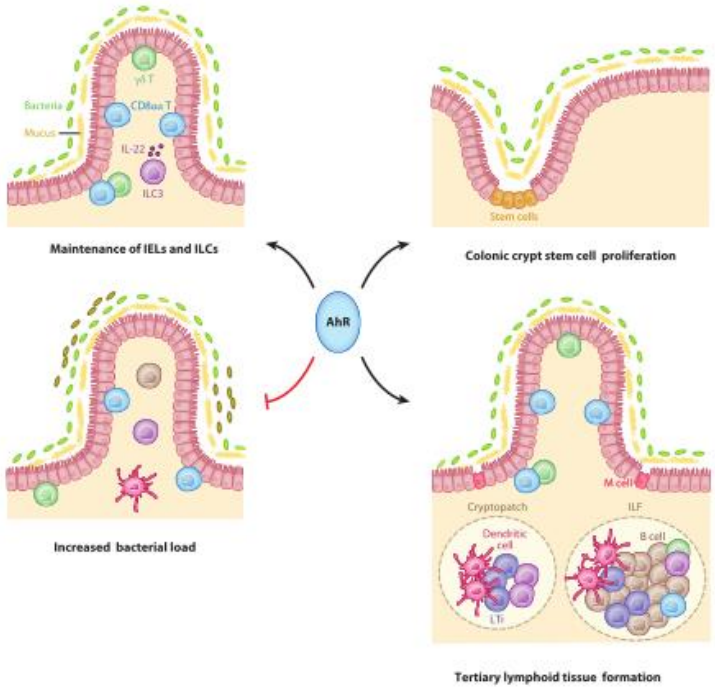


Figure 326: The role of AhR in the gut. *AhR ligand ensure the maintenance of intraepithelial lymphocytes (IELs) and innate lymphoid cells (ILC), proliferation of colonic crypt stem cells. AhR is involved in the formation of tertiary lymphoid tissues and follicles (ILFs). AhR deficiency leads to a loss of ILC3 and IELs, loss of IL-22, disruption of colonic crypt stem cell proliferation and dysregulation of intestinal bacteria (Stockinger et al. 2014).*

Role of AhR activation in IBD

Inflammatory bowel diseases (IBD) are a heterogenous group of chronic inflammatory disorders, among which Crohn's disease (CD) and ulcerative colitis (UC), characterized by the combination of genetic factors and environmental factors, among which the gut microbiota (refer to section dysbiosis in IBD). Increasing number of studies has evidenced a role of AhR in the regulation of immune response in IBD. Indeed, AhR gene and protein expressions are lower in inflamed tissues of CD patients compared to uninflamed tissues and healthy conditions (Monteleone, Rizzo, Sarra, Sica, Sileri, Biancone, MacDonald, Pallone, Monteleone, et al. 2011), suggesting the impact of AhR on the inflammatory status of, at least, CD patients. Additionally, AhR activation by ligand administration and AhR knock-down in mouse, suppressed and enhanced DSS-induced colitis respectively suggesting a beneficial effect of AhR activation in IBD (Takamura et al. 2011; Furumatsu et al. 2011). Regarding the AhR-induced protection, in a murine model of chemically induced colitis and murine cell lines, it was proposed that IECs contributes to the amelioration of disease severity by decreasing and increasing the production of pro-inflammatory cytokines (IL-6) and IgA, respectively (Benson & Shepherd 2011). Indeed, TCDD inhibited the LPS-induced stimulation of IL-6 in murine epithelial intestinal cell lines and increased the IgA production in colon (Benson & Shepherd 2011). It was then proposed that TCDD may be acting on DCs and/or IECs to produce switch factors necessary to induce IgA production by B cells in the gut and consequently drive a protective effect toward colitis via AhR activation (Benson & Shepherd 2011).

As aforementioned, an increasing number of bacterial metabolites have been identify to activate AhR, among which tryptophan- and indole-derivatives, with described protective effects against intestinal inflammation. In particular, low-Trp or sufficient-Trp diets have been associated with aggravation and protection of induced colitis in mice, respectively (Hashimoto et al. 2012). The protective effect is then proposed to be likely mediated by the microbial production of Trp-derivates, able to activate the AhR signalling pathway (Kim et al. 2010; Zelante et al. 2013). A recent work showed that the dysbiotic microbiota from mice deficient in Caspase recruitment domain 9 (CARD9), one among the numerous IBD susceptibility genes, fail to metabolize Trp with a consequent depletion in AhR ligands. This defect lead to the impaired IL-22 production and ultimately contributes to the increased susceptibility of Card9$^{-/-}$ mice to experimentally-induced colitis (Lamas et al. 2016). Accordingly, faeces of IBD

patients, especially those with CARD9 risk alleles, showed impaired AhR agonist activity, coherent with the decrease concentration of indole derivative IAA (Lamas et al. 2016). Moreover, another AhR ligand, indole-3-proprionic acid (I3P), derived from microbial metabolism of Trp, was found selectively diminished in circulating serum from human subject with active colitis (Alexeev et al. 2018). These results evidenced AhR as a possible link between the host genetics, the diet and the microbiota, in the aetiology of IBD. Coherently, other studies on animal models evidenced the importance of microbial-derived AhR ligands in the protection of intestinal inflammation. In particular, the probiotic derived molecule DHNA (1.4-Dihydroxy-2-naphthoic acid), produced by *Propionibacterium freudenreichii,* was shown to induced anti-microbial proteins RegIIIβ and RegIIIγ in the intestine of mice, with a consequent alteration of the intestinal microbiota and inhibition of experimental-induced colitis through the AhR activation (Fukumoto et al. 2014). Similar effects have been evoked also by the administration of other probiotic strains, mainly *Lactobacillus* spp., able to activate AhR pathway and improve DSS-induced colitis (Lamas et al. 2016; Takamura et al. 2011).

Overall these studies emphasize the importance of the microbial-derived AhR ligands and open the way to the identification of potential probiotic candidates in the modulation of intestinal inflammation.

Role of AhR activation in intestinal infections

One of the environmental threats to which the GI tract is exposed is pathogen infection. Together with innate and adaptive immunity, microbiota controls the resistance against colonization to maintain an ecological homeostasis (refer to section 1.4).

As previously mention, the AhR activation drives immunological responses, such as IL-22 production, which are critical for intestinal pathogen clearance. Indeed, AhR deficient mice lacking IL-22-producing ILCs in intestinal lamina propria, are more susceptible to *Citrobacter rodentium* and *Candida albicans* infections (Li et al. 2011; Lee et al. 2012; Kiss et al. 2011; Zelante et al. 2013; Qui et al. 2013). Coherently, indole 3-aldehyde, produced by various Lactobacilli from dietary Trp, activates AhR and in turn stimulates the IL-22 production by ILC3s, leading to a STAT3-dependent production of AMPs in mice with a concomitant protection against pathogen infections (Zelante et al. 2013; Qui et al. 2013). Consequently, a depleted reservoir of AhR ligand induced by a constitutive Cyp1a1 expression, results in a higher susceptibility for

Citrobacter rodentium infections (Schiering et al. 2017), further confirming the protective role of a functional AhR pathway toward intestinal infections.

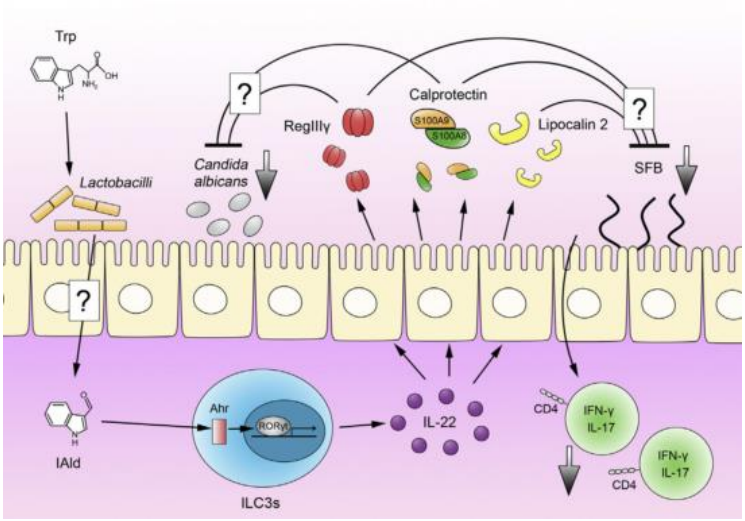


Figure 337: Mechanism of interaction between commensal bacteria and ILCs via AhR. *Lactobacilli convert the tryptophan in indolce-3-aldehyde (IAld). IAld enters the mucosa and lind to AhR in ILC3s. AhR enhance the expression of IL-22 in ILC3swhich stimulate epithelial cells to produce antimicrobial peptides. This results in an increased protection toward Candida albicans infection. AhR and retinoid-related orphan receptor g (RORgt) induce the expression of IL-22 in the small intestine, which limits expansion of SFB. Reduced levels of SFB result in a reduced amount of inflammatory inter- feron-g (IFN-g)- and IL-17-producing CD4+ T cells in the gut mucosa, protecting mice from colitis. (Behnsen et al. 2013).*

Overall these results support the protective role of AhR and a potential therapeutic application of AhR ligands in intestinal infections.

1.5.5. Cross-talk with other signalling pathways

As evidenced by the myriad of roles in physiology as well as in pathologic conditions, AhR regulates and is regulated by many signalling pathways. AhR is described to interact with steroid receptors and NFκB (Beischlag & Perdew 2005; Vogel et al. 2007).

Intensive studies described the cross-talk between AhR and **estrogen receptor** (ER), androgen receptor (AR), and thyroid hormone receptor (TR) pathways (Ghotbaddini & Powell 2015; Vrzal et al. 2017; Beischlag & Perdew 2005). Similarly to AhR, ERα and ERβ, and related steroid hormone receptors act as transcription factors dependent on the ligand binding of small hydrophobic molecules to induce their activation pathway. However, unlike AhR, ERα and ERβ, are formerly described to be activated also by a ligand-independent mechanisms (El-Tanani & Green 1997; Ding et al. 2003). The mechanisms by which PAH/HAH repress ER signalling are largely described.

Numerous evidences describes a dual role of AhR acting both as transcription inhibitor and a ubiquitin ligase to exert an anti-estrogenic effect, likely to be dependent on the estrogen concentration and the AhR agonist involved in the activation. The ligand-induced AhR:ARNT heterodimer likely binds to specific inhibitory XREs sequences (iXREs) in the promoter of ERα target genes, thus inhibiting their transcription with a consequent anti-estrogenic effect (Safe 2000). Additionally, as part of ubiquitin ligase complex, AhR exerts an antiestrogenic effect by targeting sex steroid receptors for the degradation (Ohtake et al. 2007). Conversely, the mechanisms by which ER down-regulates AhR signalling remain unclear. Evidences suggests that ERα represses TCDD-inducible Cyp1a1 and Cyp1b1 transcription through direct protein-protein interactions with the AhR:ARNT complex in the regulatory regions of these genes (Beischlag & Perdew 2005).

AhR is clearly involved in the modulation of immune response, as previously described (refer to section immunology), thus the interaction with the nuclear factor kappa B (**NFκB**) signalling could be part of the AhR mechanism for the regulation of inflammatory pathways. NFκB is a dimeric transcription factor that, upon activation, by a variety of endogenous or exogenous compounds, undergoes a rapid nuclear translocation and induces the expression of highly diverse genes, including immunoreceptors (IL-2 receptor a-chain, T cell receptor b2), cell adhesion molecules, cytokines and growth factors, chemokines, acute phase proteins, oxidative stress related enzymes and anti-apoptotic proteins (Baeuerle &

Baichwal 1997; May & Ghosh 1998). The NFκB family members can be divided into two functionally distinct classes of proteins: those synthesized as inactive precursor proteins (NFκB 1 (p105) and NFκB 2 (p100)), and those produced as transcriptionally active forms, among which RelA(p65), RelB(p66) and c-Rel have been identified in human cells (Savinova et al. 2009).

AhR and NFκB signalling pathways have been proposed to interact through a physical association of their respective critical components, RelA/RelB and AhR, thus associated with a mutual functional modulation of the genes controlled by both AhR and NFκB (Tian et al. 1999; Vogel et al. 2007). By a direct interaction, RelA and AhR proteins have been described to functionally cooperate to bind to NFκB elements and induce the expression of the oncogene c-myc in breast cancer cell line. Additionally, the AhR/RelB dimer is capable of binding to DNA response elements, including XRE as well as NFκB binding sites, supporting the activation of target genes of the AhR as well as NFκB pathway (Vogel et al. 2007). Moreover, activators of NFκB signalling, such as lipopolysaccharide and proinflammatory cytokines (IL-1β; tumor necrosis factor (TNF)-α), have been described to suppresses the TCDD-induced CYP1A1 expression (Ke et al. 2001). In addition, through the cooperation with NFκB signalling, AhR has been evidenced to promote cell proliferation in human mammary cells but, discordantly with other cell lines (Kim et al. 2000; Ito et al. 2004).

Altogether the presented scenario, even if not exhaustive, shows the complexity of the interactions established by AhR and let the open question on lots of physiological roles, not yet described, in which the receptor could take part.

Chapter 2. Rationale and Objectives

The mammalian gastrointestinal tract (GI) is colonized by a complex microbial community, referred as gut microbiota, in close contact with the intestinal mucosa. In the intestine, the co-evolution of the mammalian host with its microbial community, shapes the intestinal cells to sense a variety of bacterial metabolites, establishing a fine regulated human-microbiota cross-talk for the maintenance of the intestinal homeostasis. While the GI tract offers nutrients for the sustainment of the microbial community, metabolites derived from commensal bacteria have been evidenced for having a strong impact on mucosal homeostasis. Some microbial metabolite, such as short-chain fatty acids (SCFA), from the fermentation of dietary fibers, and indoles, from the degradation of dietary tryptophan, target several host receptors. In turn, these receptors induce specific signalling pathways that modify host gene expression and collectively impact both the host metabolism and immune responses. Additionally, the host-microbiota interaction is fundamental in the maintenance of a physical and ecological barrier against pathogen infection and of the intestinal homeostasis.

Recently the ligand-activated transcription factor, aryl hydrocarbon receptor (AhR), has gained increasing interest as a critical regulator of both adaptive immunity and mucosal barrier functions (Stockinger et al. 2014). Expressed in a wide variety of cell types (among which ILCs, IECs, macrophage, dendritic cells), especially in the barrier organs, the precise role of AhR emerged over the years, and it is nowadays described as a sensor of environmental stimuli, among which the bacterial metabolites. Indeed, extensive studies revealed that AhR is capable of binding not only dioxins but also a wide range of structurally diverse molecules that originate from the diet, the environment, the microbiota or produced endogenously by the host itself (Barouki et al. 2007; Hubbard, Murray, Bisson, et al. 2015). To date, the bacterial-induced activation of AhR is described to derived from a limited number of bacterial species, including some Lactobacilli (Firmicutes) and Proteobacteria (Sonowal et al. 2017; Zelante et al. 2013), and been principally mediated by the production of indole or indole-derived molecules, from tryptophan catabolism (Lamas et al. 2018). The microbial-induced AhR activation has been evidenced as protective toward pathogen infection or experimentally-induced colitis, suggesting a pivotal role of AhR in the maintenance of gut homeostasis and a potential therapeutic target. Few non-indoles bacterial metabolites have been describes, such as the

DHNA (1,4-Dihydroxy-2-naphthoic acid) produced by the probiotic *Propionibacterium freudenreichii* (Fukumoto et al. 2014) as well as some bacterial pigments producing by the pathogens *Pseudomonas aeruginosa* and *Mycobacterium tuberculosis* (Moura-Alves et al. 2014). These evidences provide an additional element to the described promiscuity of AhR toward binding molecules of different origins. However, despite the structurally pleiotropic nature of AhR ligands and the huge metabolic potential of the gut microbiota, the identified commensal-derived molecules are still limited. Therefore, it emerged the need for an extended knowledge on microbial-derived AhR activators, to decipher the host-microbiota interaction via AhR, and better described the actors in the modulation of host epithelial homeostasis.

In this context, the present work, funded by MetaCardis European project (FP7 - HEALTH-F4-2012-305312), aims to identify novel key microbial metabolites able to activate AhR in IECs and potentially involved in intestinal homeostasis. To attain this scientific object, the project is structure in two parts:

- Paper I: The screening of a commensal bacterial collection to identify new microbial actors in the AhR activation in IECs and the consequent identification of the cellular signalling involved in the activation of AhR by the bacterial-produced butyrate;
- Paper II: The identification of a new bacteria genus activating AhR signalling, *Bifidobacterium* spp.

This project lead to the identification of two distinct mechanisms for bacterial activation of AhR in IECs that made the object of the two publications presented below.

Chapter 3. Results

3.1. *Paper I: Identification of the novel role of butyrate as AhR ligand in human intestinal epithelial cells.*

This first article covers the major part of the presented work on the identification of microbial metabolites activating AhR signalling in human intestinal epithelial cells (IECs). To identify activator microbial species and strains, we screened a collection of commensal bacteria on an AhR reporter system expressed in a human intestinal epithelial cell line (HT-29). Then, to characterize the molecules involved in the activation, we quantified the main organic acids described to be produced by gut microbiota. Consequently, we confirmed the identified bacterial activators on another intestinal epithelial cell line (Caco-2) and we described the mechanism for the bacterial-induced AhR activation.

This study reveals that (i) some bacterial strains activate AhR signaling pathway through the production of butyrate and (ii) we highlighted for the first time that butyrate acts as a novel ligand of AhR that stimulates the translocation of the receptor in the nucleus and induces the downstream expression of AhR-regulated genes *CYP1A1* and *AhRR*, together with *AHR* gene and protein expression.

This project was funded by MetaCardis European project (FP7 - HEALTH-F4-2012-305312).

Identification of the novel role of butyrate as AhR ligand in human intestinal epithelial cells.

Ludovica Marinelli[1,2], Camille Martin-Gallausiaux[1,2], Jean-Marie Bourhis[3], Fabienne Béguet-Crespel[1], Hervé M. Blottière[1, 4] and Nicolas Lapaque[1*].

1. Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, France

2. Sorbonne Universités, UPMC Univ Paris 06, IFD, 4 place Jussieu, 75252, Paris, cedex 05, France

3. University Grenoble Alpes, CNRS, CEA, IBS, F-38000, Grenoble, France

4. MetaGenoPolis, INRA, Université Paris-Saclay, 78350 Jouy en Josas, France

* Corresponding author:

Dr Nicolas Lapaque

INRA-MICALIS UMR1319, Bat 442, Domaine de Vilvert 78350, Jouy-en-Josas, France.

e-mail: nicolas.lapaque@inra.fr

Abstract

The ligand activated transcription factor, aryl hydrocarbon receptor (AhR) emerged as a critical regulator of immune and metabolic processes in the gastrointestinal tract. In the gut, a main source of AhR ligands derives from commensal bacteria. However, many of the reported microbiota-derived ligands have been restricted to indolyl metabolites. Here, by screening commensal bacteria supernatants on an AhR reporter system expressed in human intestinal epithelial cell line (IEC), we found that the short chain fatty acid (SCFA), butyrate, induced AhR activity and the transcription of AhR-dependent genes in IECs. We showed that AhR ligand antagonists reduced the effects of butyrate on IEC suggesting that butyrate could act as a ligand of AhR, which was supported by the nuclear translocation of AhR induced by butyrate and *in silico* structural modelling. In conclusion, our findings suggest that (i) butyrate activates AhR pathway and AhR-dependent genes in human intestinal epithelial cell-lines (ii) butyrate is a potential ligand for AhR which is an original mechanism of gene regulation by SCFA.

Introduction

The mammalian gastrointestinal (GI) tract is colonised by a complex microbial community, referred as gut microbiota. It is well established that host-commensal bacteria crosstalk provides numerous functions for the overall host wellbeing, through the production of microbial metabolites. The host-microbiota interaction is particularly substantial for mucosal barrier functions as well as the development and maintenance of the mucosal immune system[1]. Metabolites derived from commensal bacteria described to strongly impact mucosal homeostasis include the short-chain fatty acids (SCFA), originated from the fermentation of dietary fibres, and indoles, from the degradation of dietary tryptophan. The targeted host receptors of these bacterial products include the cell-surface G-protein-coupled receptors GPR41, GPR43, and GPR109A and nuclear receptors such as aryl hydrocarbon receptor (AhR), pregnane X receptor (PXR), and farnesoid X receptor (FXR). The receptor-metabolite interactions induce signalling pathways that modulate host gene expression and collectively impact on host metabolism and immune responses[1].

Recently AhR, a ligand activated transcription factor, has gained considerable attention as a crucial modulator of mucosal immune and metabolic processes, especially in the context of diet and microbiota crosstalk with the host[2-4]. AhR is a member of the basic helix-loop-helix PER-ARNT-SIM (bHLH-PAS) family, initially identified as a hepatic intracellular protein that bounds with high affinity the environmental halogenated contaminant 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). However, extensive studies revealed that AhR interacts with a wide range of structurally diverse molecules that originate from the diet, the environment, the microbiota or are produced endogenously by the host itself[5,6]. Many of the microbiota-derived AhR ligands result from tryptophan catabolism including indole, indole-3-acetic acid, indole-3-aldehyde[7,8]. Upon ligand-binding, cytoplasmic AhR translocates in the nucleus, dimerizes with AhR nuclear translocator (ARNT) and initiates the transcription of target genes with promoters containing a xenobiotic-response element (XRE) sequence. Genes such as cytochrome P450 family 1A1 (*CYP1A1*) and the repressor AHRR are regulated by AhR activation[9,10].

AhR activation has been largely reported to be implicated in colonic stem cells proliferation, epithelial barrier functions, maintenance of intraepithelial lymphocytes (IEL), innate lymphoid cells (ILC) and FOXP3 regulatory T cells (Treg)[2-4]. Interestingly, AhR signalling and known AhR ligands are low in inflammatory bowel diseases (IBD) patients, highlighting the clinical relevance of the AhR pathway in these disease[11,12]. AhR activation by ligand administration and AhR knock-down in mouse suppressed and enhanced DSS-induced colitis, respectively suggesting a beneficial effect of AhR activation in IBD[13,14]. An increasing number of bacterial metabolites have been shown to drive the AhR activation, with described protective effects against intestinal inflammation and pathogens colonisations, suggesting a possible role of this signalling pathway in the intestinal homeostasis[7]. This hypothesis has been further illustrated by a recent study showing that one IBD-associated single

nucleotide polymorphism (SNP) within the *CARD9* gene, affects microbiota composition thus altering the production of bacterial AhR ligands and consequently intestinal inflammation[11].

The relevant role of AhR in the maintenance of homeostasis at mucosal surfaces has been largely focused on immune cells from the *lamina propria*, and the impact of AhR activation in intestinal epithelial cells (IEC) has only been starting to be unravelled. Emerging evidences highlight that AhR activation impacts also IECs responses contributing to intestinal homeostasis. AhR activation in IECs protects mice from colitis by enhancing IECs barrier functions via the increase of the IL10 receptor expression and the enhancement of tight junctions integrity through the regulation of *Notch1*[15,16]. Importantly, the excess of AhR ligands in the gut has been shown to be detrimental to the host. In this context, a recent study in mice model nicely demonstrates that IECs play the role of gatekeeper *via* the expression of the AhR-regulated gene *Cyp1a1* encoding cytochrome P450 1A1, an enzyme involved in AhR ligands clearance[17,18]. These studies highlighted the importance of AhR ligands in modulating host gut immune homeostasis and prompted us to identify new microbiota-derived activators of the AhR pathway in IECs. We thus tested the bacterial supernatants of over 100 bacterial species of the human microbiota on an AhR reporter system in human intestinal cell lines and found that butyrate-producing bacteria activate the AhR-dependant response. We confirmed the butyrate-activating role on AhR reporter system and, at the transcriptional level, on AhR-dependent genes in Caco-2 and HT-29 cell lines. Antagonists blocking the binding of AhR ligand impaired the butyrate-induced activation of AhR reporter system and the up-regulation of *CYP1A1* gene expression highlighting for the first time that butyrate could act as a ligand of AhR.

Results

Metabolites derived from commensal bacteria enhanced AhR activity

In the gut, Aryl hydrocarbon receptor (AhR) ligands derive from diverse origins that include the intestinal microbiota as one of the main source. To decipher which bacteria from human gut activate the AhR pathway, we performed a screening of commensal bacteria on a human intestinal epithelial cell line (HT-29-AhR) stably expressing an AhR-dependant reporter system. The AhR reporter system contains three copies of the DNA-binding domain motif recognized by AhR (xenobiotic-responsive element; XRE) driving the transcription of the luciferase reporter gene. Since AhR activators produced by bacteria are secreted in the intestinal lumen, we performed the screening of bacterial culture supernatants[8]. The screening included 132 bacterial strains, belonging to the major phyla of the human intestinal microbiota (Actinobacteria, Bacteroidetes, Firmicutes, Fusobacteria and Verrucomicrobia) (Supplementary Fig S1, Supplementary Table S1) and grown under appropriate bacterial growth conditions (Supplementary Table S2). When possible, we cultured the bacteria in different media, to avoid biased results due to the rich composition of culture media, at least for almost all the bacteria activating the AhR reporter system. In our experimental set-up, AhR activation was detected in HT-29 cells challenged with some supernatants derived from Proteobacteria, Firmicutes, Fusobacteria and few Actinobacteria (Fig 1).

Butyrate activates AhR pathway in intestinal epithelial cell lines

Despite the huge literature on AhR ligands, only few are commensal-derived molecules. Amongst them, indole and other tryptophan derivatives are produced by a wide range of bacteria including *Lactobacillus* (Firmicutes) and Proteobacteria[19,20]. In our screening we confirmed the AhR activation induced by some lactobacilli as well as by members of the Proteobacteria, probably due to the production of indole derivatives. Interestingly, among the bacteria not reported or not predicted to be indole producers, we identified some members of genera *Ruminoclostridium* and *Roseburia,* belonging to the Firmicutes phylum. These genera, together with other AhR-activating bacteria predicted to produce indole (*Clostridium* and *Lachnoclostridium*) share the ability to degrade diet-derived fibres, leading to the production of short-chain fatty acids (SCFAs)[21]. We thus hypothesized that SCFAs concentration in the supernatants of these bacteria could explain the activation of the AhR pathway. We therefore quantified the concentrations of SCFAs in some bacterial supernatants (Supplementary Table S1 and Fig 2A). A principal component analysis (PCA) of the complete dataset revealed a strong positive correlation between AhR activation and the production of butyrate (Fig 2B). This analysis was confirmed by the positive relationship between butyrate concentration and AhR activity (rho=0.4966) obtained by Spearman correlation (Fig 2C).

To reinforce experimentally the observed correlations, we tested pure SCFAs at different physiological concentrations found in the intestine (ranging from 0.125mM to 8mM) on HT-29-AhR reporter cells (Fig 2D, Supplementary Fig S2A)[22,23]. Among the tested SCFAs, we confirmed that butyrate was able to induce AhR activation in a dose-dependent manner, in HT-29-AhR cell line as well as in another intestinal reporter cell line, Caco2-AhR in a dose-dependent manner (Fig 2D, Supplementary Fig S2B). Propionate, another abundant SCFA produced by bacteria, was also able to activate AhR in a dose-dependent manner in HT-29-AhR cells at a concentration starting at 1mM. Interestingly, we also observed that valerate and the branched chain fatty acids iso-valerate, both described to influence epithelial physiology, were also able to activate AhR pathway in HT-29-AhR cell line at similar concentrations[24,25] Iso-butyrate activated AhR only at 8mM a concentration rarely reached in the intestine, while acetate, the most abundant SCFA produced by commensal bacteria did not shown any activation (Supplementary Fig S2A). The SCFA concentrations inducing AhR activity were consistent with the final SCFAs concentrations on bacterial supernatants measured in the screen thus reinforcing our hypothesis (Supplementary Table S1).

Moreover, we showed by qRT-PCR that AhR-regulated genes, *CYP1A1*, *AHR* and *AHRR*, were highly up-regulated by butyrate both in HT-29 and Caco2 cells confirming that this SCFA activated AhR pathway (Fig 3A, Supplementary Fig S3). We confirmed the butyrate-driven up-regulation of AHR at the protein level by Western-Blot analysis in HT29 (Fig 3B).

AhR activation by butyrate is independent of the SCFA receptors GPR41, GPR43, GPR109a and the SCFAs transporter MCT-1.

Butyrate, like other SCFAs, activates eukaryotic cells through two main mechanisms: activation of specific G-protein coupled receptors (GPCR: GPR41, GPR43, GPR109a) and inhibition of histone deacetylases (HDAC)[26-30]. Considering that the three G-protein coupled receptors are expressed in HT-29 and Caco-2 cells (Martin-Gallausiaux *et al.* submitted), we firstly tested the hypothesis that the butyrate-induced AhR activation could be mediated by GPRs signalling. Thus, we stimulated HT29-AhR cells with known GPRs agonists, targeting GPR41, GPR43 and GPR109a. For each GPR, two agonists were tested (GPR41: 4-CMTB and Tiglic acid; GPR43: AR420626 and 1-MCPC; GPR109a: Niacine and MK1903). Interestingly, none of the tested agonists induced AhR activation in luciferase reporter system suggesting that these GPRs may not be involved in the butyrate-induced activation in HT-29-AhR and Caco2-AhR cell lines (Fig 4A and Supplementary Fig S4). GPR41, GPR109a are both Gαi coupled receptors whereas GPR43 is a Gαi and Gαq coupled receptor. To further confirm our observation, we used an inhibitor of the Gαi pathway inhibitor: the pertussis toxin (Ptx) in HT-29-AhR (Fig 4B). No impact on the butyrate-induced AhR activation was detected in cells when the $G\alpha_i$ subunit was inhibited, further confirming that these GPRs were not involved in the observed AhR activation.

Butyrate uptake by the monocarboxylate transporter MCT-1 has been described as crucial for the GPR-independent regulation of a wide range of genes by butyrate[31]. To assess if MCT-1 participated in the activation of AhR by butyrate, a well-described inhibitor of MCT-1 transporter, pCMB, was tested on HT-29-AhR cells (Fig 4C). The inhibition of MCT-1 transporter did not affect the ability of butyrate to activate AhR reporter system, suggesting that MCT-1 was not involved in this process.

AhR activation by butyrate is independent of its role as inhibitor of HDAC

SCFAs, *via* their ability to inhibit lysine and histone deacetylases (HDAC), are potent modulators of histones and transcription factors acetylation that are well-documented regulatory mechanisms of gene transcription[26-28]. A recent study showed that SCFAs and other HDAC inhibitors (HDACi) enhanced AhR-induced genes such as *CYP1A1 via* the increase level of histone acetylation[32]. These results prompted us to investigate the role of HDACi in AhR activation. To assess if butyrate impacted AhR activity through its HDACi property, we tested three HDACi targeting a wide range of HDAC, trichostatin A (TSA), Vorinostat (SAHA) belonging to the hydroxamic acids family, structurally and metabolically unrelated to SCFAs and sodium valproate (VAP) belonging to the fatty acid family[33]. TSA, SAHA and VAP did not reproduce the butyrate-induced AhR activation in both HT-29-AhR and Caco2-AhR reporter system, suggesting that the inhibition of HDAC was probably not involved in the cellular mechanism observed (Fig 5A, Supplementary Fig S5A). Interestingly, we could reproduce the activation of *CYP1A1* by TSA as observed by Jin *et al* (Fig 5B)[32]. Altogether, our results suggested that HDACi property of SCFAs was not involved in the butyrate-dependent AhR activation although HDACi inhibition might participate in the regulation of AhR-induced genes independently of AhR activation.

Butyrate acts as an AhR ligand

As showed previously, we ruled out the two main mechanisms described for butyrate to activate transcription factors, namely its HDACi property and the activation pathway *via* GPRs. Two groups described an original mechanism of gene regulation where butyrate acts as a ligand of the transcription factor PPARγ[34,35]. As AhR is a ligand binding transcription factors, we hypothesized that butyrate could act as an AhR ligand similarly to what was described for PPARγ. To demonstrate the potential role of butyrate as an AhR ligand, we tested the butyrate-induced AhR responses in the presence of well-characterized inhibitors of the AhR pathway. We thus incubate butyrate with two antagonists of the AhR-ligand binding (CH-223191 and GNF-351) and one HSP90 inhibitor, described to avoid the downstream AhR-XRE binding, by blocking the release of HSP90 chaperon from the AhR complex (epigallocatechine-3-gallate, EGCG)[36-39]. The three tested inhibitors significantly decreased the butyrate-induced activation of AhR in HT-29 and Caco-2 cell lines (Fig

6A, Supplementary Fig S5B). In addition, we confirmed in HT-29 cells that CH-223191 and GNF-351 were able to block *CYP1A1* expression induced by butyrate (Fig 6B). Altogether, these results suggested that the activation of AhR pathway by butyrate was due to a ligand-dependent AhR-XRE interaction, indicating a possible role of butyrate as a direct AhR ligand.

It is well established that ligand binding triggers the accessibility of the nuclear localisation signal on the AhR N-terminus that consequently initiates the AhR nuclear translocation event [40]. Thus, additional evidence for the human AhR agonist potential of butyrate was obtained by performing nuclear translocation assay. Sub-cellular localisation of AhR in HT-29 incubated with butyrate (1, 3 and 6h) and TCDD (1h) was assessed by immunoblotting assay. We showed an accumulation of AhR proteins in the nucleus upon treatment with butyrate starting at 3h and with a peak at 6h, consistent with the action of an AhR agonist such as TCDD (Fig 7).

Altogether, our results suggested that butyrate activation of the AhR reporter system relied on AhR translocation to the nucleus and consequent DNA-binding and that butyrate might act as a ligand of AhR.

In silico modelling of butyrate interaction with AhR

Numerous AhR modulators, binding in large central pocket of the AhR PAS-B domain have been extensively studied using a combination of homology modelling and docking stimulations (for review see[41]. To support our results suggesting that butyrate acts as an AhR ligand, we performed a molecular docking simulation of the butyrate/AhR binding process. We first generate a homology model of human AhR ligand binding domain (AhR PAS-B) using HIF-2α PAS-B domain as template similarly to Bisson *et al.*[42,43]. The AhR PAS-B homology domain was obtained with RaptorX modelling software and the binding cavity was defined using HOLLOW and Carver softwares (Fig 8A). Then, the model was minimized to avoid clashes between side chains. From the best scoring solutions cluster using HADDOCK, we observed different orientations of the butyrate inside the AhR PAS-B cavity. Based on four different docking experiments, two main orientations were found for butyrate in the pocket of human AhR PAS-B domain. Interactions between butyrate and AhR are dominated by polar contacts with the side chains of Q383 and S365, for the first orientation (Fig 8B) and with side chains of Q383 and H291, for the second one (Fig 8C). Interestingly, by comparing our results with published docking analyses of other ligands on human AhR PAS-B, the first proposed orientation for butyrate docking shared the same interacting side chains (Q383 and S365) with the docking of FICZ, proposed by Bisson *et al.*[42]. Overall these molecular docking analyses are coherent with our findings of butyrate as a direct modulator of AhR by ligand binding.

Discussion

AhR is a ligand-activated transcription factor that is crucial for intestinal homeostasis by repressing inflammation and by maintaining the epithelial barrier in the gastrointestinal tract. AhR activation regulates both adaptive immunity and mucosal barrier functions[3,44]. AhR has been shown to regulate a wide range of immune cell populations including RORγt+ innate lymphoid cell-3 (ILC3), T helper (Th)17/22 cells, intraepithelial γδT cells, regulatory T cells and antigen presenting cells [45,46]. The AhR-dependent *IL22* expression by ILC3 regulates the release of antimicrobial peptides such as RegIIIγ and the expression of tight junctions molecules in IECs thus reinforcing the barrier and defence functions[19,47]. Recently, emerging evidences highlight that AhR directly impacts IECs leading to increased barrier functions and the regulation of excess AhR ligands in the gut[17,18]. The gastrointestinal tract is a rich source of AhR ligands, which have been shown to induce AhR dependent responses and to protect the gut upon infection or induced colitis. Despite the huge literature on the structurally pleotropic nature of AhR ligands, only few are commensal-derived molecules. Amongst them, indoles and other tryptophan derivatives are produced by a variety of bacteria including some Lactobacilli (Firmicutes) and Proteobacteria[19,20]. Interestingly, non-commensal bacteria have been reported to produce non-indole AhR ligands such as the phenazine derivative from *Mycobacteria* and the 1,4-dihydroxy-2-naphthoic acid from the probiotic *Propionibacterium freudenreichii*[47,48]. Given the abundance and extensive metabolic capacity of gut microbiome, it is likely that metabolites apart from tryptophan derivatives are present in high concentration to stimulate AhR pathway.

By screening commensal bacteria supernatants, we identified butyrate as a potent activator of AhR pathway and AhR-dependent genes. Butyrate is a short chain fatty acid (SCFA) that derived from bacterial fermentation of dietary fibers. Butyrate *via* its role as inhibitor of lysine/histone deacetylase (K/HDACi), influences expression of a large variety of host genes in the colon, including some encoding for immune proteins[49]. Previous studies described butyrate as a regulator of AhR-dependent genes thought its role as histone deacetylase inhibitors (HDACi)[32,50]. However, these studies assessed the butyrate effect through induction of AhR transcriptional targets not directly on the AhR receptor. Our results are in accordance with these studies as we showed that HDACi and butyrate upregulate *CYP1A1* expression. However, we demonstrated that butyrate does not impact AhR-dependent gene expression solely by its HDACi properties. Indeed by using an AhR reporter system in HT-29 and Caco2 cells, we showed that butyrate, and not HDAC inhibitors, activates AhR signalling pathway.

To decipher the mechanism of butyrate-dependent AhR activation, we investigated the implication of butyrate specific G-protein coupled receptors (GPR41, GPR109a and GPR43) and main transporter, monocarboxylate transporter 1 (MCT1)[51-53]. By using agonists and inhibitors, we showed that AhR activation by butyrate was not mediated by those GPRs or MCT1. Many studies evidenced that the predominant biological activities of the AHR are through ligand binding. Despite

the AhR association with xenobiotic compounds, structurally diverse metabolites from the diet, bacteria or produced by the host have been reported as capable of binding to human AhR[6,54]. Our experimental results suggest that butyrate acts as an AhR ligand leading to the activation of the AhR pathway. Moreover, structural modelling of the binding of butyrate to human AhR PAS-B support this hypothesis as we observed two main orientation of butyrate in the ligand binding cavity of AhR. Interestingly, one of this orientation showed polar contact of butyrate with the same side chains of PAS-B reported for the binding of FICZ to human AhR supporting our hypothesis that butyrate is a ligand of AhR[42]. This is an original mechanism of butyrate-dependent modulation of host gene expression, which has only been reported for another ligand-dependent transcription factor, PPARγ. Indeed, two groups reported that PPARγ-dependent genes regulation by butyrate is independent of its HDACi properties and is mediated through a direct binding of the SFCA to this transcription factor[34,35]. The strong impact of butyrate on mucosal immune homeostasis has been largely documented in mice and *in vitro* models. In human, studies have shown lower concentration of SCFA and butyrate-producing bacteria (*e.g. Faecalibaterium* and *Clostridium*) in the gut of inflammatory bowel diseases (IBD) patients suggesting a relevant role of butyrate in intestinal health[55-57]. Butyrate and other SCFAs elicit most of their biological activities by binding to GPRs and by acting as HDAC inhibitors[49]. Our results demonstrate an additional mechanism where butyrate binds to AhR thus activating its signalling pathway leading to the expression of AhR-dependent genes. Interestingly, SCFAs such as butyrate exhibit overlapping activities with AhR ligands on intestinal homeostasis and in IBD[58]. AhR ligands regulate epithelial IL-10 receptor α subunit (IL-10RA) expression that damped colitis by promoting epithelial wound healing[59]. Similarly, butyrate promotes epithelial barrier formation through IL-10RA induction on IECs[60]. Hence, it is possible that AhR mediates the butyrate-induced IL-10RA upregulation. Considering the high quantities of SCFAs in the gut, it is likely that butyrate possibly in synergy with other bacterial-derived AhR ligands have a role in the physiological functions of AhR.

In conclusion, we show that butyrate stimulates AhR-dependent genes through a direct AhR activation and probably in complement to its HDACi property in human intestinal cell. Our results suggest that butyrate acts as a ligand of AhR which is, to our knowledge, an original mechanism only been reported for another ligand-binding transcription factor PPARγ[35].

Materials and Methods

Cell Culture of human colonic cell lines

The human epithelial cell lines HT-29 and Caco-2 were obtained from the American Type Culture Collection (ATCC, Rockville, MD). HT-29 were grown in RPMI 1640 GlutaMAX™ and Caco-2 in DMEM GlutaMAX™ medium supplemented with 10% and 20% of heat-inactivated fetal bovine serum (FBS, Lonza), respectively. Both media were supplemented with 50 IU/mL penicillin, 50 µg/mL streptomycin and 10%, 100mM Hepes, 10mM nonessential amino acids. HT-29 and Caco-2 were grown at 37°C in a humidified 5% and 10% CO2 atmosphere, respectively. All culture media and supplements were supplied by Gibco (ThermoFisher). Mycoplasma contamination was regularly tested using MycoAlert (Lonza) and PlasmoTest (Invivogen).

Production of Stable AhR-luciferase Reporter Cell-Lines

pGL4.43[luc2P/XRE/Hygro] (Promega) was used to establish HT-29-AhR and Caco2-AhR reporter cell-lines by electroporation using the Nucleofector® device (Lonza) according to the manufacturer's recommendations. Stable AhR reporter cell lines were selected using Hygromycin (600 µg/ml for HT-29 and 200 µg/ml for the Caco2 cell line, InvivoGen) and validated using TCDD at 10nM final concentration.

Culture of commensal bacteria, preparation of supernatants and SCFA concentration assessment

132 human intestinal commensal bacterial strains (106 different species) from the in-house INRA-Micalis collection or from DSMZ were grown. Anaerobic culture conditions were done accordingly to the Hungate method[61]. Screened strains, corresponding growth media, optical densities (OD600), SCFA concentrations are listed in Appendix Table S1 and composition of home-made growth media is listed in Supplementary Table S2. Bacterial cultures were cultured to reach the maximum OD. Bacterial supernatants were harvest after centrifugation at 5,000 × g for 10 min and filtered on a 0.22µm PES filters and stored at -80°C. Quality controls were performed using Gram staining method, aerobic growth test and fresh observation on microscope. Non-inoculated bacteria culture medium served as control. Concentrations of SCFAs produced by cultured bacteria were measured by HPLC and gas chromatography as described by[62,63].

Luciferase Reporter and Cell Viability Assays

For the bacterial screening, HT-29-AhR cells were seeded at 3×10^4 cells per well in white 96-well plates (Corning). After 24 h from seeding, cells were stimulated during 24 hours with 10 µL of bacterial supernatant or un-inoculated media in a total culture-volume of 100 µL per well (10% vol/vol). The screening was performed in triplicates and for almost all the samples, experiments were performed at least with two biological replicates. Additionally, when possible, some strains were grown in different bacterial media. For testing the effect of reagents on AhR activity on HT-29 and

Caco-2 cells, 24h after seeding the culture media was replaced with a non-FBS-supplemented RPMI or DMEM. The cells were then stimulated with 10µL of reagents diluted in non-FBS conditions in a total culture-volume of 100 µL per well (10% vol/vol). Follow-up experiments were performed in triplicates and repeated at least three times. Luciferase activity was quantified as relative luminescence units (RLU) using a microplate reader (infinite® 200 plate reader, TECAN) and the Neolite™ (PerkinElmer) Luciferase Assay System according to the manufacturer's instructions. The AhR activation was normalised on non-inoculated bacterial media or untreated/vehicle-treated cells for bacterial supernatants and tested reagents, respectively. The results were expressed as luciferase fold change. Cell viability was assessed by MTS measurement using the CellTiter 96 Aqueous One solution (Promega) according to the manufacturer's recommendations.

Reagents

All agonists, antagonists and drugs tested were dissolved in a proper vehicle (DMSO, glycerol, water, PBS or ethanol) following the manufacturer's recommendations. The final concentration used for vehicles had not detectable effect on metabolic activity of the cells. Sodium salts of tested SCFAs were from Sigma and used in a range of concentration from 0.125 to 8mM. AhR agonist: 2,3,7,8-Tetrachlorodibenzodioxin (TCDD 10nM, Sigma). GPRs agonists: GPR41: 4-chloro-α-(1-methylethyl)-N-2-thiazolylbenzeneacetamide (4-CMTB 1µM, Tocris) and Tiglic acid (1mM, Sigma); GPR43: N-(2,5-Dichlorophenyl)-4-(furan-2-yl)-2-methyl-5-oxo-1,4,5,6,7,8-hexahydro-quinoline-3-carboxamide (AR420626 1µM, Cayman) and 1-methylcyclopropane carboxylate (MCPC 1mM, Sigma); GPR109a: Niacine (1mM, Sigma) and (4aR,5aR)-4,4a,5,5a-Tetrahydro-1H-cyclopropa[4,5]cyclopenta[1,2]pyrazole-3-carboxylic acid (MK1903 10nM, Tocris). Pertussis toxin (Ptx at 0.02µg/ml, Sigma) was used as $G\alpha_i$-subunit inhibitor. MCT1 inhibitor used was p-Chloromercuribenzoate acid (pCMB 100µM, Sigma). HDAC inhibitors: Trichostatin A (TSA 0.1µM, Sigma), vorinostat (SAHA 5µM, Sigma) and valproic acid (VPA 5mM, Sigma). AhR antagonists: CH-22319 (1µM, Millipore/Calbiochem), GNF-351 (1µM, Millipore/Calbiochem), (-)Epigallocathechin gallate (20µM, EGCG, Sigma).

Real-Time PCR

Cell lines were seeded in 12-well culture plates at densities of 0.5×10^6 cells per well. The cells were seeded in FBS-supplemented media then, after 24h, the media was replaced with a non-FBS-supplemented and cells incubated during 24h before stimulation. After stimulation time of 6h, total RNA was extracted using RNeasy mini-Kit (Qiagen) according to manufacturer's recommendations. cDNA was synthesized from 2 µg of RNA using the High-Capacity cDNA Archive Kit (Applied Biosystems). qPCRs were carried out using an StepOne (Applied Biosystems) thermal cycler in a reaction volume of 20 µL with Taqman gene expression assay probes: *AHR*: Hs00169233_m1; *AHRR*: Hs01005075_m1; *CYP1A1*: Hs01054796_g1; *β-ACTIN*: Hs99999903_m1. *AHR*, *CYP1A1* and *AHRR* expression relative to control expression was determined by the $2^{-\Delta\Delta Ct}$ method using *β-*

actin as control gene. Data are means ± SEM of at least three distinct experiments, performed in triplicate.

Cytoplasmic and nuclear protein extraction

HT-29 cells were seeded at densities of 0.5×10^6 cells per well in 12-well-plates. 24h after seeding the media was replaced with a non-FBS-supplemented RPMI and cells incubated during 24h prior to stimulation. When nuclear extracts were not needed, cells were washed twice and lysed in buffer (1% NP40, 150mM NaCl, 50mM Tris-HCL pH8, 5mM EDTA, 1 x Complete Protease Inhibitor Cocktail (Roche)). Nucleus were eliminated by centrifugation for 10 minutes 4°C at 17500g. For compartments separation, nuclear and cytoplasmic extracts were prepared with NE-PER Nuclear and Cytoplasmic Extraction Reagent Kit (ThermoFisher) according to the manufacture instructions. CER I and NER buffers were supplemented with protease inhibitor cocktail (cOmplete™ ULTRA Tablets, Mini, *EASY*pack Protease Inhibitor Cocktail, Sigma) prior to use.

Western Blot analysis

Protein extracts were run in 10% SDS-PAGE gels and transferred onto PVDF membranes by liquid transfer (Transfer buffer: 192mM Glycine, 25mM TrisBase, 20% methanol) at 200mA during 90minutes. Membranes were blocked overnight in TBST+4% BSA (Sigma). Primary antibodies were incubated overnight at 4°C: anti-AhR (1:500, mouse mAb, clone RTP1, ThermoFisher), anti-Lamin A/C (1:2000, mouse mAb, Cell Signaling), anti-GAPDH (1:4000, mouse mAb, Santa Cruz), anti-Lamp1 (1:2000, mouse mAb, H4A3 from the Developmental Studies Hybridoma Bank (DSHB), H4A3 was deposited to the DSHB by August, J.T./Hildreth J.E.K. (DSHB hybridoma product H4A3). Secondary mouse horseradish peroxidase-coupled antibody (DAKO) was successively incubated at room temperature for 2h before detection with the Clarity Western ECL Substrate using the Chemidoc MP System (Bio-Rad). Quantifications were performed using the image Lab software (Bio-Rad). AhR nuclear protein levels were normalised to Lamin A/C protein levels. Lamp1 and GAPDH were used as purification controls for the cytoplasmic proteins.

Modeling of butyrate binding to AhR

For modelling the structure of the complex between AhR and butyrate, HADDOCK software was used[64,65]. HADDOCK is a highly successful modelling approach that makes use of structural knowledge when available to drive the docking procedure. In this case the crystal structure of the heterodimeric HIF-2α:ARNT complex (PDB code 4ZP4)[66] was used as template in order to dock butyrate to human AhR similarly to Bisson *et al.*, using RaptorX[42,43,67,68]. The cavity of AhR PAS-B was defined by CARVER and HOLLOW[69,70]. The figures were generated with PyMOL Molecular Graphic System, version 1.8 Schrödinger, LLC.

Statistical analysis

Presented results are representative of at least three independent experiments. The PCA analysis and Speaman correlation were performed using R and RStudio software. Graphics were produced using Prism GraphPad software. The data distribution was tested using D'Agostino-Person omnibus normality test. Normally distributed data was checked using two-sided t test, otherwise, non parametric Wilcoxon signed-rank or Mann-Whitney tests were performed according to the data set. In all tests, ns: $P>0.05$, *: $P\leq0.05$, **: $P\leq0.01$, ***: $P\leq0.001$, ****: $P<0.0001$.

Acknowledgments

The authors are grateful to Dr Marion Espeli (INSERM UMR-S996) and to the member of the team Functionality of the Intestinal Ecosystem (Micalis-INRA Jouy-en-Josas) for helpful discussions. The authors are grateful to Agnès David and Catherine Michel for the quantification of SCFA in bacterial cultures. This work was supported by the Institut National de la Recherche Agronomique (INRA), by grants funded by EU-FP7 METACARDIS (HEALTH-F4-2012-305312).

Author Contributions

Conceived and designed the experiments: LM, NL, HMB; performed most of the experiments: LM; performed some experiments: CMG, FBC, JMB, NL; analysed the data: LM, NL; wrote the paper: LM, NL; edited and revised the manuscript: CMG, JMB, HMB.

Conflict of interest

The authors disclose no conflict of interest.

Figure legends

Figure 1:

Screening of in-house strain collection of commensal bacteria on HT-29-AhR cell line. HT-29-AhR reporter cells were incubated with bacterial supernatants or relative non-inoculated bacterial media for 24h (10% vol/vol). AhR activation was measured by luciferase activity and expressed as fold increase (± SD from triplicates) toward its control (non-inoculated bacterial media), sorted by Phyla.

Figure 2:

Butyrate activates AhR pathway. (A) HPLC quantification of butyrate produced by tested bacterial supernatants, sorted by Phyla (N.D.: Not Determined). (B) Principal Component Analysis (PCA) of quantified SCFAs and AhR activation for the entire data set. The axes PC1 and PC2 accounted for

34,83% and 55.58% of total data variation, respectively. (C) Spearman correlation shows a positive relationship (rho=0.4966) between AhR activity and butyrate concentration classified by rank values. (D) HT29-AhR reporter cells were incubated with a range of concentration of butyrate rising (0.125mM to 8mM). Data are expressed as luciferase fold (± SEM) of at least three independent experiments, normalised on untreated cells. ns: P>0.05, *: P≤0.05, **: P≤0.01, ***: P≤0.001, ****: P<0.0001.

Figure 3:

Butyrate activates the expression of AhR-regulated genes. (A) The expression of AhR-regulated genes: *AHR*, *CYP1A1* and *AHRR* on HT-29 cells treated with butyrate 2mM or TCDD 10nM during 6h was determined by qRT-PCR. *AHR*, *CYP1A1* and *AHRR* relative expression to control is determined by the $2^{-\Delta\Delta Ct}$ method using *β-actin* for normalisation. (B) Up-regulation of AhR protein level by butyrate. HT-29 cells were incubated with butyrate 2mM or TCDD 10nM for 24h. Total cytoplasmic extracts were blotted (Western Blot) for AhR (left panel) GAPDH was used as loading control. Relative quantification is expressed as fold-change to un-stimulated control of AhR protein normalised on GAPDH level (right panel).

Data are expressed as means ± SEM of at least three distinct experiments, performed in triplicate. ns: P>0.05, *: P≤0.05, **: P≤0.01, ***: P≤0.001, ****: P<0.0001.

Figure 4:

Mechanism involved in AhR activation by butyrate is independent of GPR and MCT-1. (A) HT-29-AhR reporter cells were stimulated for 24h with GPR agonists. GPR41 (orange): 1-MCPC (1mM), AR-420626 (1µM); GPR43 (blue): Tiglic acid (1mM), MK-1903 (1µm); GPR109a (green): Niacine (1mM), 4-CMTB (1µM). (B) HT-29-AhR reporter cells were incubated with $G\alpha_i$-subunit inhibitor, Pertussis toxin (Ptx, 0.2ug/ml) prior a 24h butyrate stimulation (2mM). Vehicle (Glycerol) was used as control. (C) HT29-AhR reporter cells were incubated during 24h with a MCT-1 inhibitor, *p*-chloromercuribenzoic acid (pCMB, 10uM), prior stimulation with butyrate (2mM).

AhR activation was measured by luciferase activity and expressed as fold increase means (± SEM) of at least three independent experiments, normalised on un-treated cells. ns: P>0.05, *: P≤0.05, **: P≤0.01, ***: P≤0.001, ****: P<0.0001.

Figure 5:

Impact of HDAC inhibitors on AhR reporter system and *CYP1A1*. (A) HT29-AhR reporter cells were stimulated with HDAC inhibitors for 24h. Trichostatin A (TSA, 0.1mM), Vorinostat (SAHA 5µM), Valproic acid (VAP 5mM), butyrate (But, 2mM) and TCDD (10nM). (B) Relative expression to control of *CYP1A1* on HT-29 cells treated with butyrate 2mM, TCDD 10nM or TSA 0.1mM during

6h. *CYP1A1* expression induced by drugs is expressed as relative to control expression and is determined by the $2^{-\Delta\Delta Ct}$ method using *β-ACTIN* as control gene.

AhR activation was measured by luciferase activity and expressed as fold increase means (± SEM) of at least three independent experiments, normalised on un-treated cells. ns: P>0.05, *: P≤0.05, **: P≤0.01, ***: P≤0.001, ****: P<0.0001

Figure 6:

AhR antagonists inhibit AhR activation by butyrate. (A) HT-29-AhR reporter cells were incubated with AhR ligand antagonists (CH-223191 1µM and GNF-351 1µM) and an HSP90 inhibitor ((-)-Epigallocatechin-3-gallate, EGCG, 20µM) prior stimulation with butyrate (But, 2mM) and TCDD (10nM) for 24h. (B) Relative expression of *CYP1A1* on HT-29 cells treated during 6h with 2mM butyrate or TCDD 10nM in presence or absence of AhR antagonists. *CYP1A1* expression is normalised to control expression (unstimulated cells) and is determined by the $2^{-\Delta\Delta Ct}$ method using *β-actin* as control gene.

Data are means ± SEM of at least three distinct experiments, performed in triplicate. ns: P>0.05, *: P≤0.05, **: P≤0.01, ***: P≤0.001, ****: P<0.0001.

Figure 7:

Butyrate stimulates AhR protein translocation in the nucleus. HT-29 cells were incubated during 1h, 3h or 6h in presence of 2mM butyrate or 1h with TCDD. Nuclear and cytoplasmic extracts were blotted (Western Blot) for AhR protein expression. GAPDH and Lamp were used as control for cytoplasmic extract preparation; Lamin A/C as control for nuclear extract preparation (left panel). Relative quantification of AhR protein level in nuclear extracts from least six independent experiments is expressed as fold-change (± SEM) to un-stimulated cells (control) after normalisation on Lamin A/C level. ns: P>0.05, *: P≤0.05, **: P≤0.01, ***: P≤0.001, ****: P<0.0001.

Figure 8:

Modelling of butyrate binding to the ligand binding pocket of human AhR. (A) Model of PAS-B domain of human AhR represented in cartoon coloured in wheat, the central cavity is represented as a red surface. Model of PAS-B domain of human AhR using 4ZP4 Crystal Structure of the Heterodimeric HIF-2a:ARNT complex obtain using RaptorX. The cavity has been defined by CARVER and HOLLOW. (B) Q383/S365 orientation. The residues Q383 and S365 are displayed as sticks and coloured by atom type with carbon in wheat. Butyrate is displayed as sticks and coloured by atom type with carbon in magenta. Hydrogen bonds are represented by white dashed lines. (C) Q383.H291 orientation. The residues Q383 and H291 are displayed as sticks and coloured by atom type with carbon in wheat. Butyrate is displayed as sticks and coloured by atom type with carbon in green. Hydrogen bonds are represented by white dashed lines. The figures were generated by PyMol.

Supplementary Figures

Supplementary Figure S1:

Screening of in-house strain collection of commensal bacteria on HT-29-AhR cell line. AhR activation is expressed as the fold increase (± SD) toward its control (non-inoculated bacterial media), sorted by bacterial strain. Bacteria are sorted by response in decreasing order and grouped by Phyla (Actinobacteria in green, Bacteroidetes in orange, Firmicutes in purple, Fusobacteria in light blue, Proteobacteria in light green, Verrucomicrobia in pink). HT29-AhR cells were exposed to bacterial supernatants or relative non-inoculated bacterial media for 24h (10%vol/vol).

Supplementary Figure S2:

Effect of short chain fatty acids (SCFAs) on HT29-AhR and Caco2-AhR reporter cell lines.
(A) HT-29-AhR reporter cells were incubated with SCFAs at concentrations rising from 0.125mM to 8mM. (B) Caco2-AhR reporter cells were stimulated with of butyrate at concentration rising from 0.125mM to 8mM.
AhR activation was measured by luciferase activity and expressed as fold increase means (± SEM) of at least three independent experiments, normalised on un-treated cells. ns: $P>0.05$, *: $P\leq0.05$, **: $P\leq0.01$, ***: $P\leq0.001$, ****: $P<0.0001$

Supplementary Figure S3:

Relative expression to control of *AHR*, *CYP1A1* and *AHRR* genes on Caco-2 cells treated with butyrate (2mM) or TCDD (10nM) during 6h. *AHR*, *CYP1A1* and *AHRR* expression are expressed relative to control (un-stimulated cells) are determined by the $2^{-\Delta\Delta Ct}$ method using *β-ACTIN* as control gene. Data are expressed as means ± SEM of at least three distinct experiments, performed in triplicate.

Supplementary Figure S4:

Effect of GPR agonists on Caco2-AhR cells. Caco2-AhR reporter cells were stimulated for 24h with GPR agonists. GPR41 (orange): 1-MCPC (1mM), AR-420626 (1µM); GPR43 (blue): Tiglic acid (1mM), MK-1903 (1µm); GPR109a (green): Niacine (1mM), 4-CMTB (1µM).
AhR activation was measured by luciferase activity and expressed as fold increase means (± SEM) of at least three independent experiments, normalised on un-treated cells. ns: $P>0.05$, *: $P\leq0.05$, **: $P\leq0.01$, ***: $P\leq0.001$, ****: $P<0.0001$

Supplementary Figure S5:

Effect of HDAC inhibitors on Caco2-AhR cells. (A) Impact of HDAC inhibitors on Caco2-AhR reporter system. Caco2-AhR reporter cells were stimulated with HDAC inhibitors for 24h. Trichostatin A (TSA, 0.1mM), Vorinostat (SAHA 5 uM), Valproic acid (VAP 5 mM), butyrate (But,

2mM) and TCDD (10nM). AhR activation was measured by luciferase activity and expressed as fold increase means (± SEM) of at least three independent experiments, normalised on un-treated cells. (B) AhR antagonists inhibit AhR activation by butyrate in Caco2 cells. Caco2-AhR reporter cells were incubated with AhR ligand antagonists (CH-223191 1uM and GNF-351 1uM) and an HSP90 inhibitor ((-)-Epigallocatechin-3-gallate, EGCG, 20uM) prior stimulation with butyrate (But, 2mM) and TCDD (10nM) for 24h.

Activation of the AhR reporter system was measured by luciferase activity, normalised on un-treated cells and expressed as fold increase means (± SEM) of at least three independent experiments. ns: $P>0.05$, *: $P\leq0.05$, **: $P\leq0.01$, ***: $P\leq0.001$, ****: $P<0.0001$

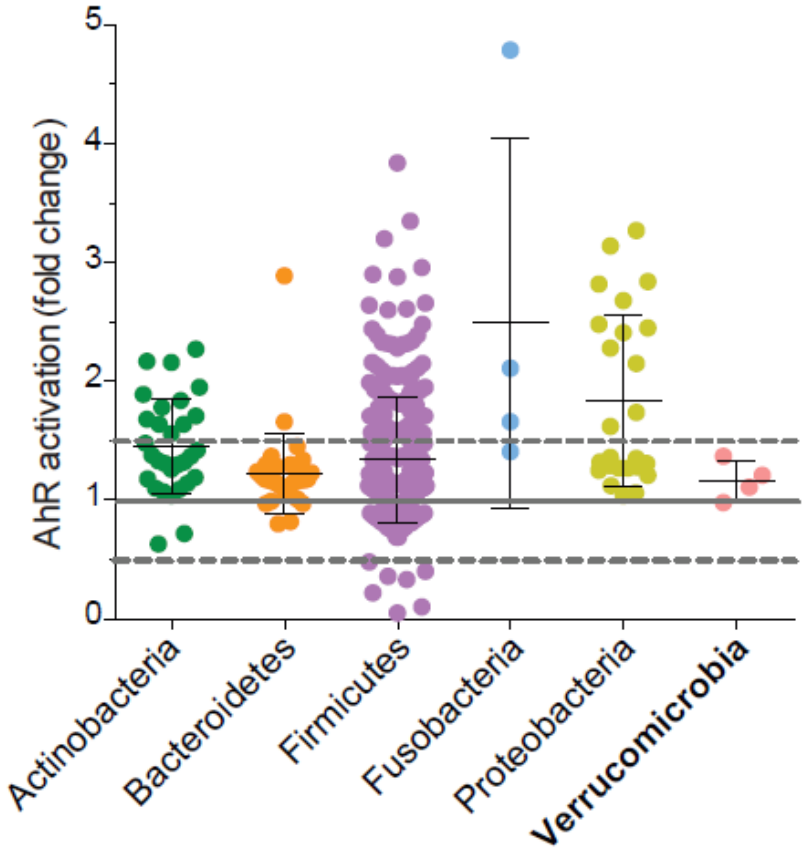
Supplementary Tables

Supplementary Table S1:

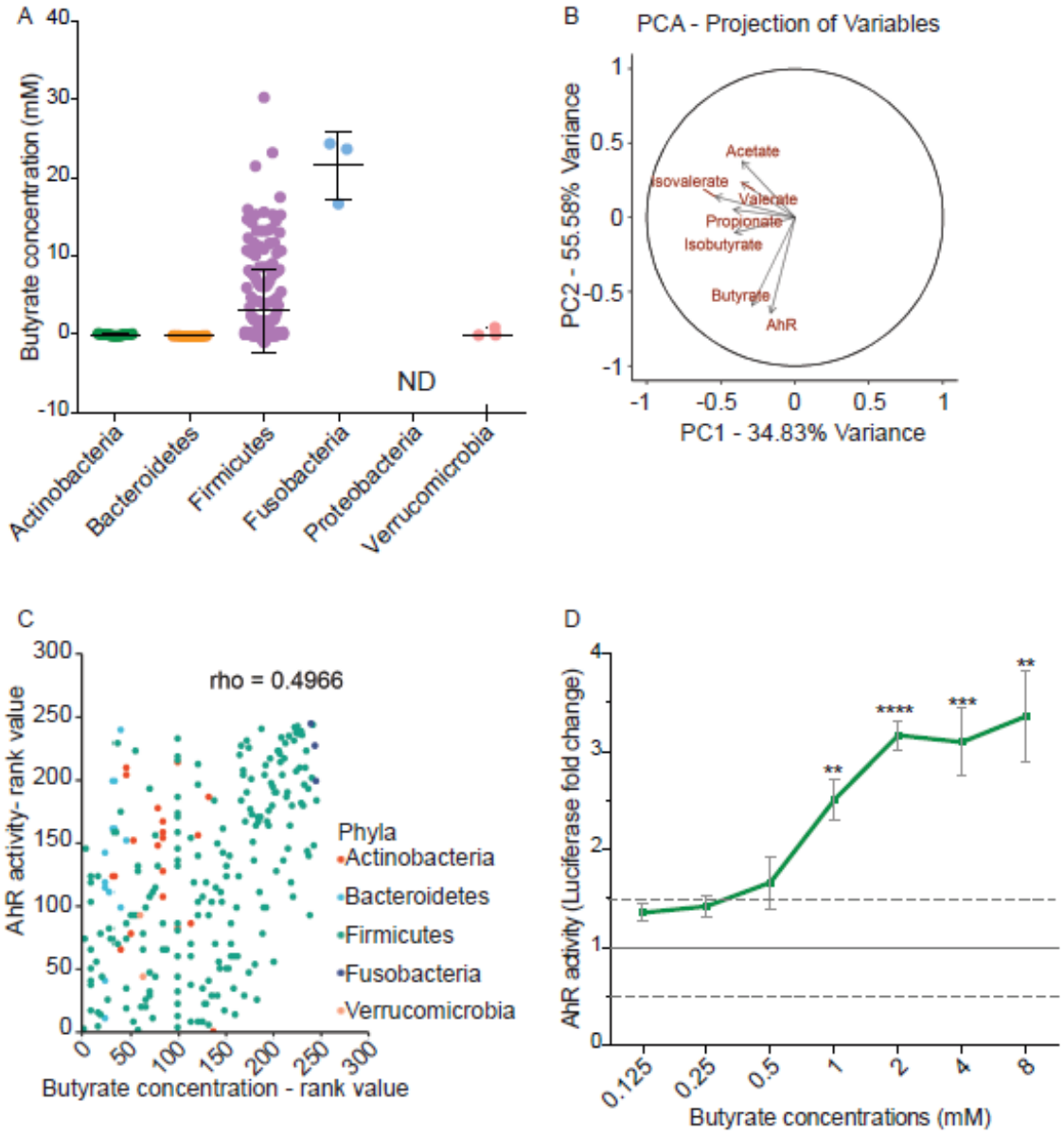
List of all tested commensal bacterial strains, bacterial growth (measured by optical density 600, OD600), pH of conditioned medium, SCFA concentrations and AhR activity (as fold increase as compared to control medium)

Supplementary Table S2:

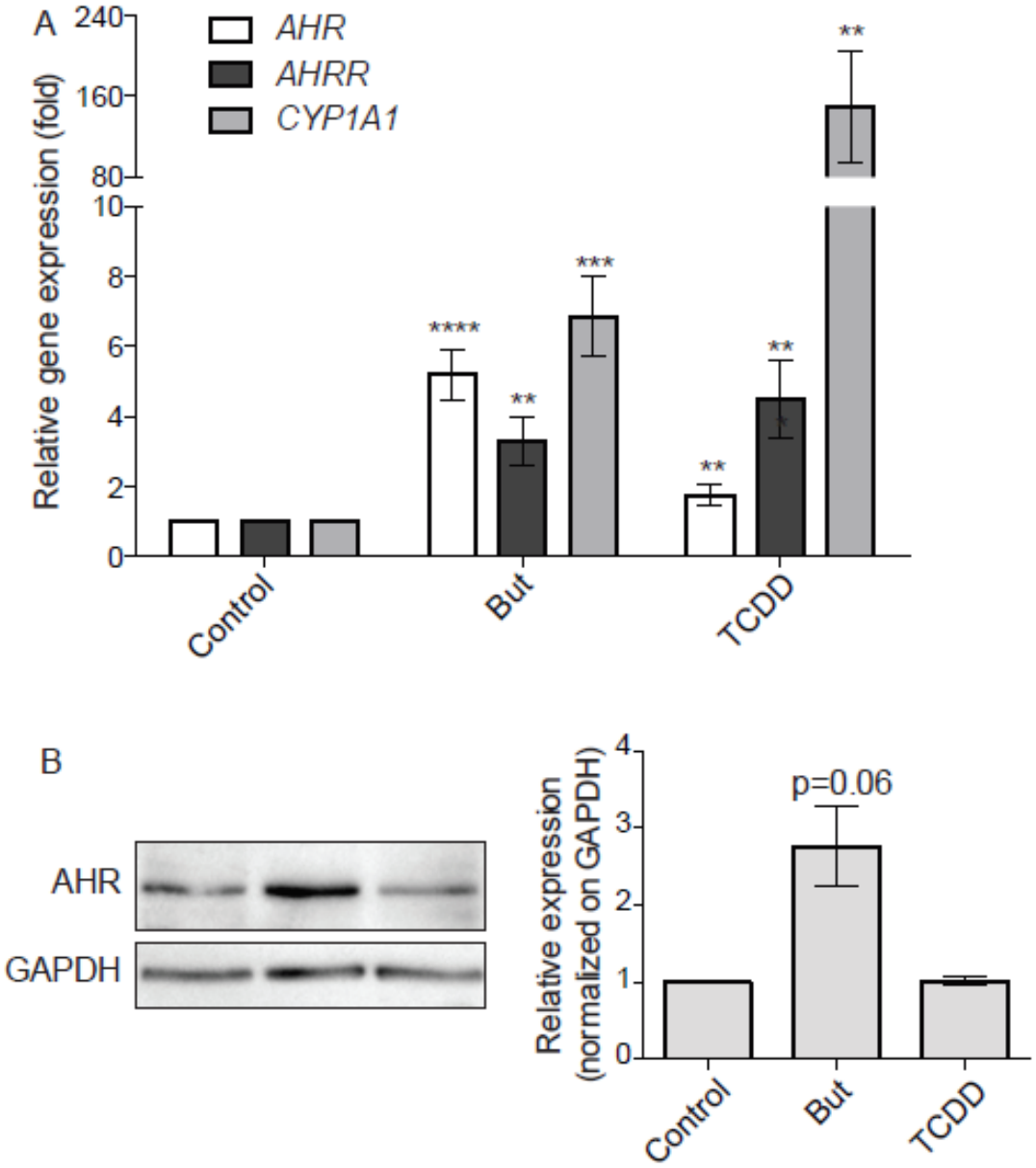
Composition of bacterial media used in the study.



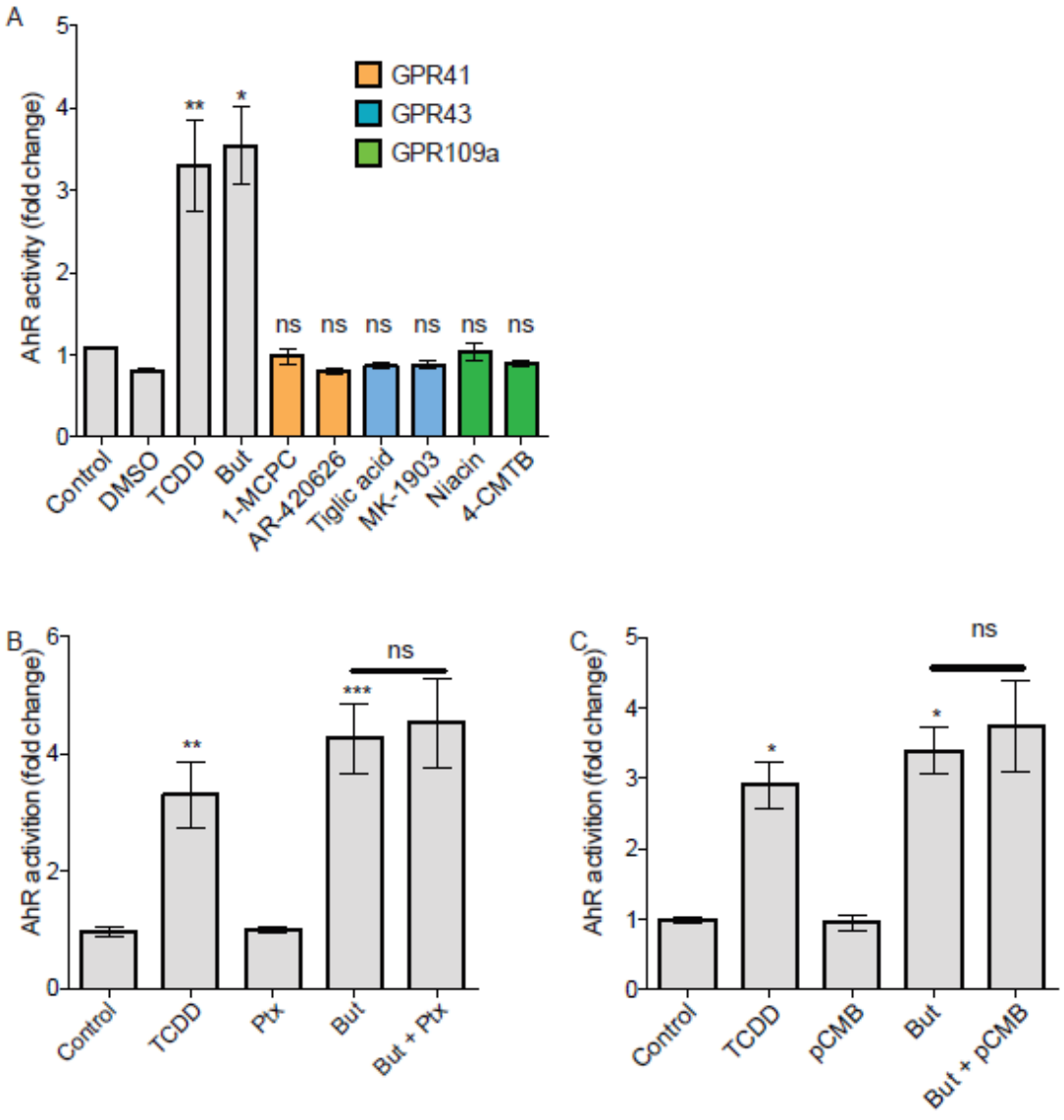
Marinelli *et al.*
Figure 1



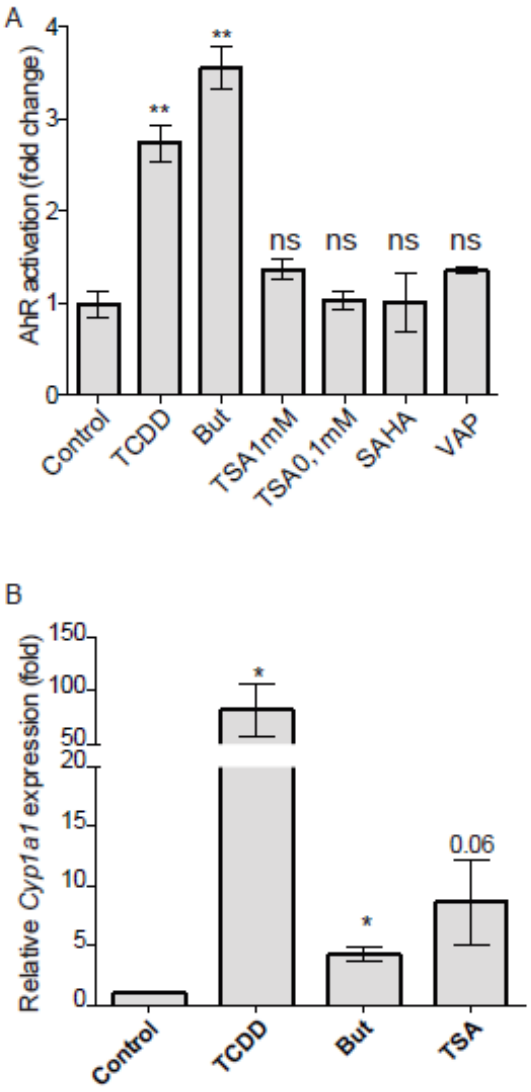
Marinelli *et al.*
Figure 2



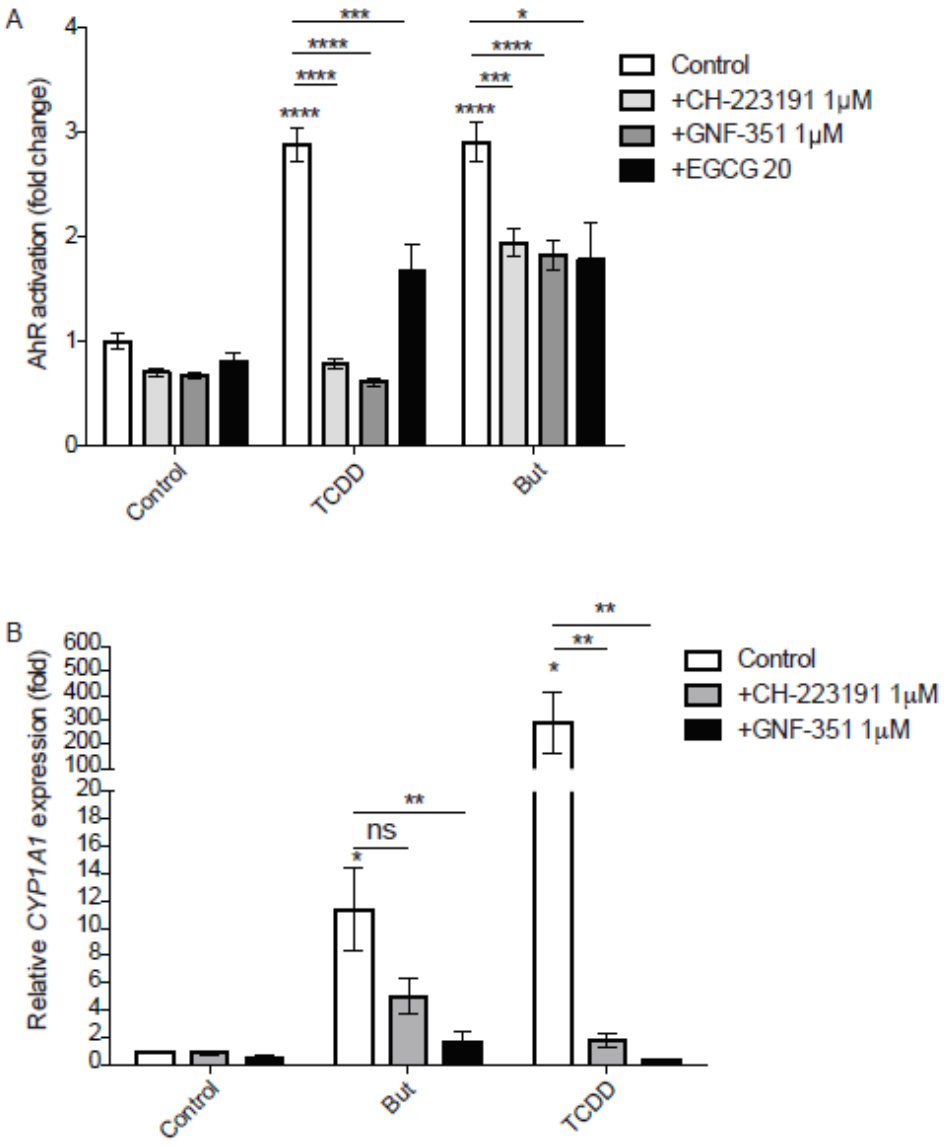
Marinelli *et al.*
Figure 3



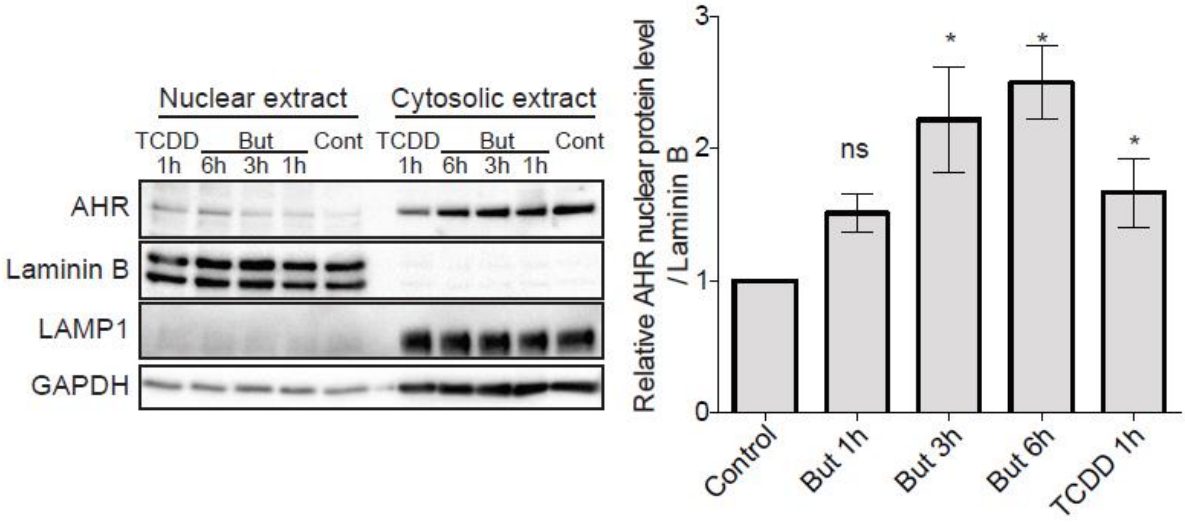
Marinelli *et al.*
Figure 4



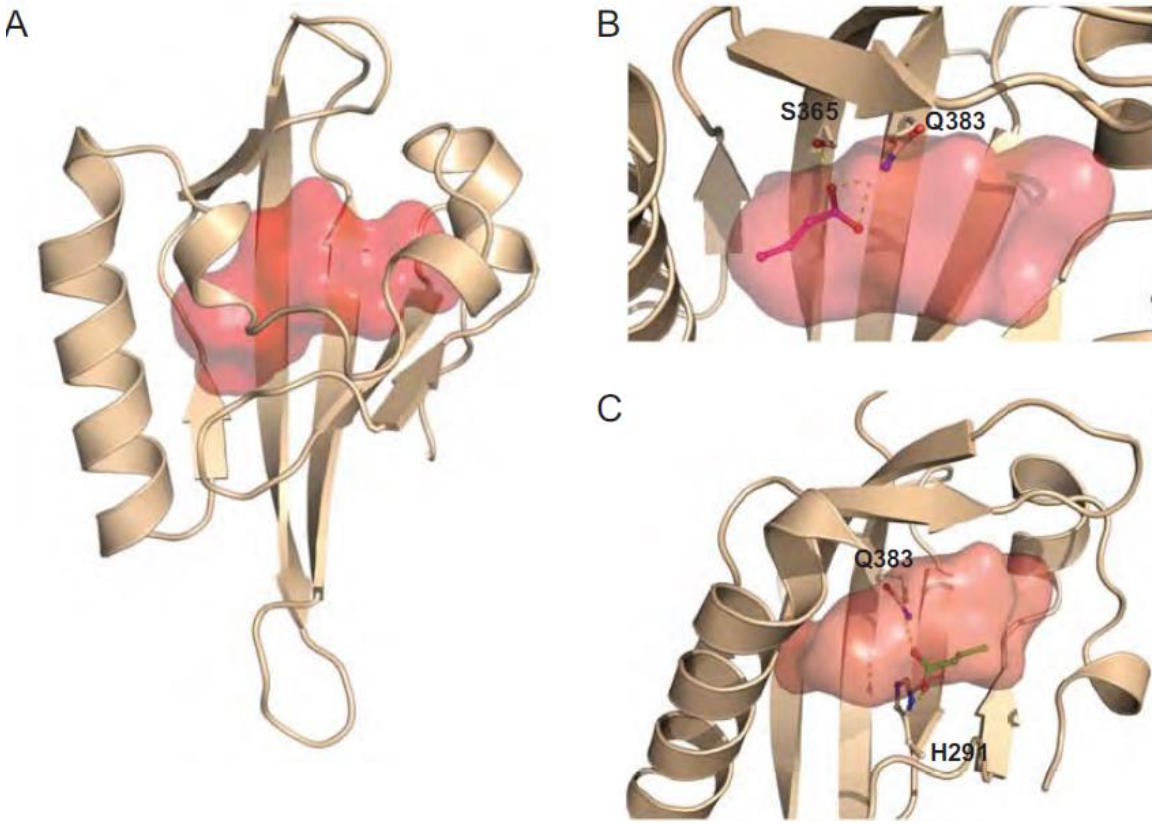
Marinelli *et al.*
Figure 5



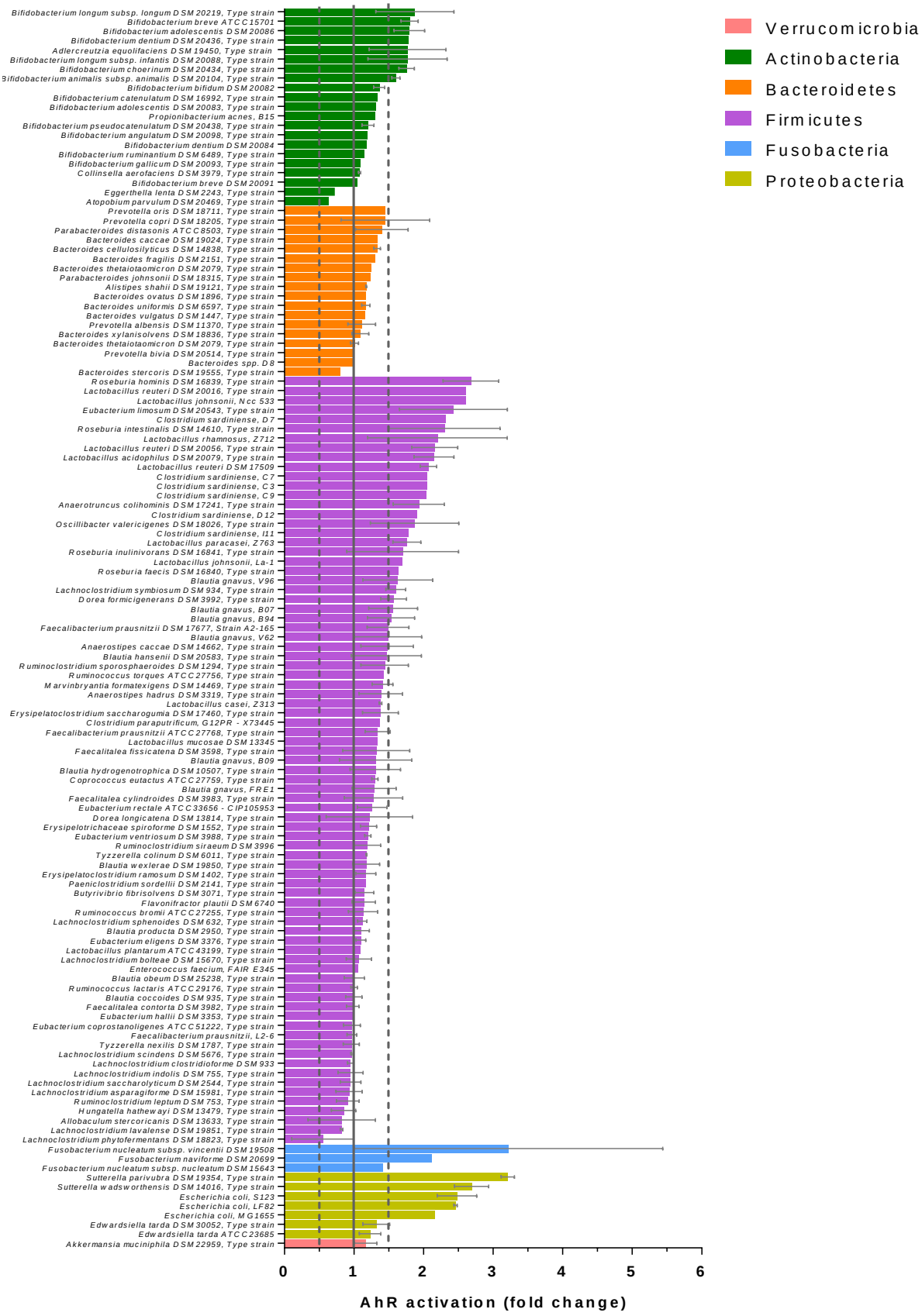
Marinelli *et al.*
Figure 6



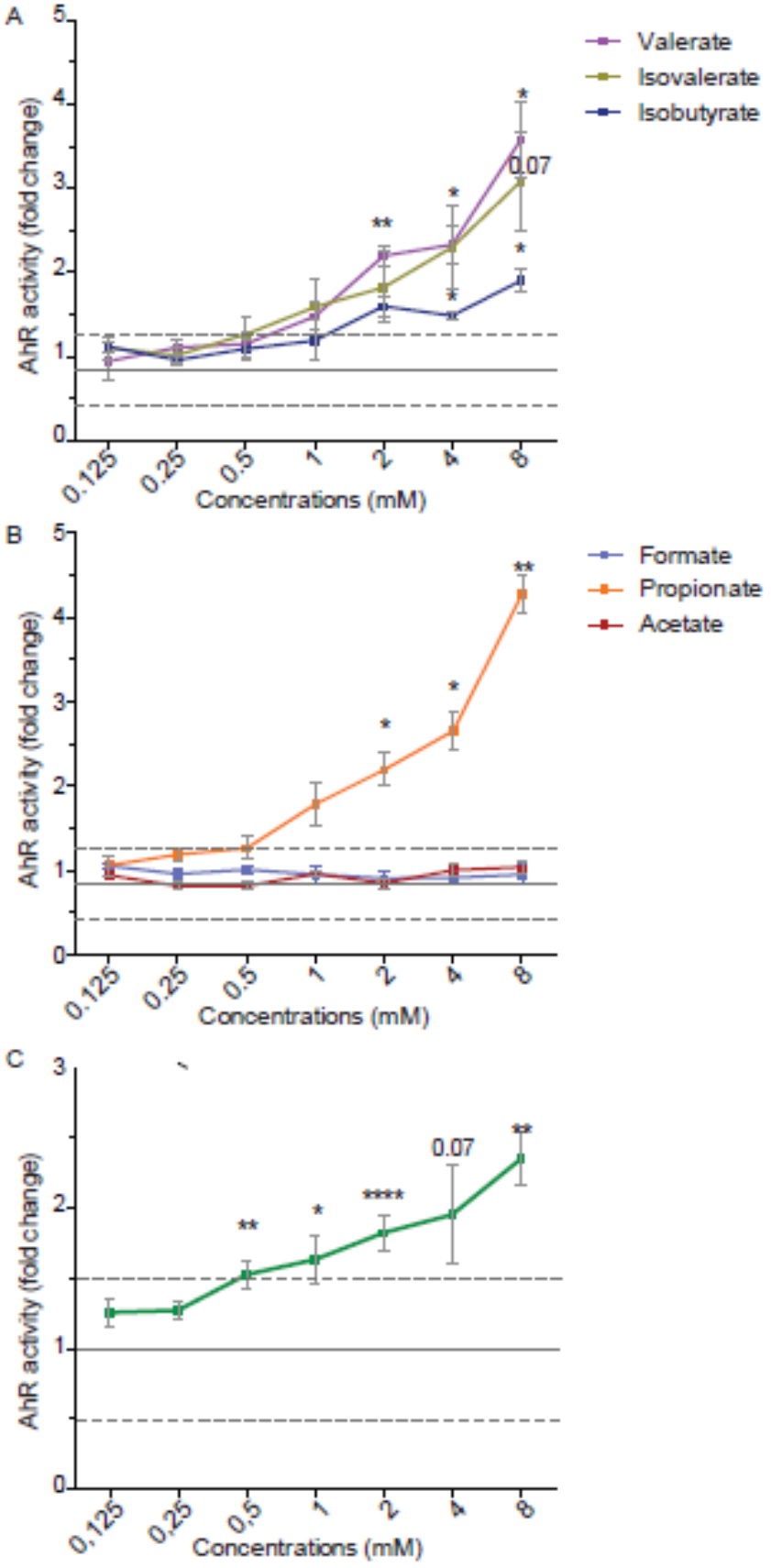
Marinelli *et al.*
Figure 7



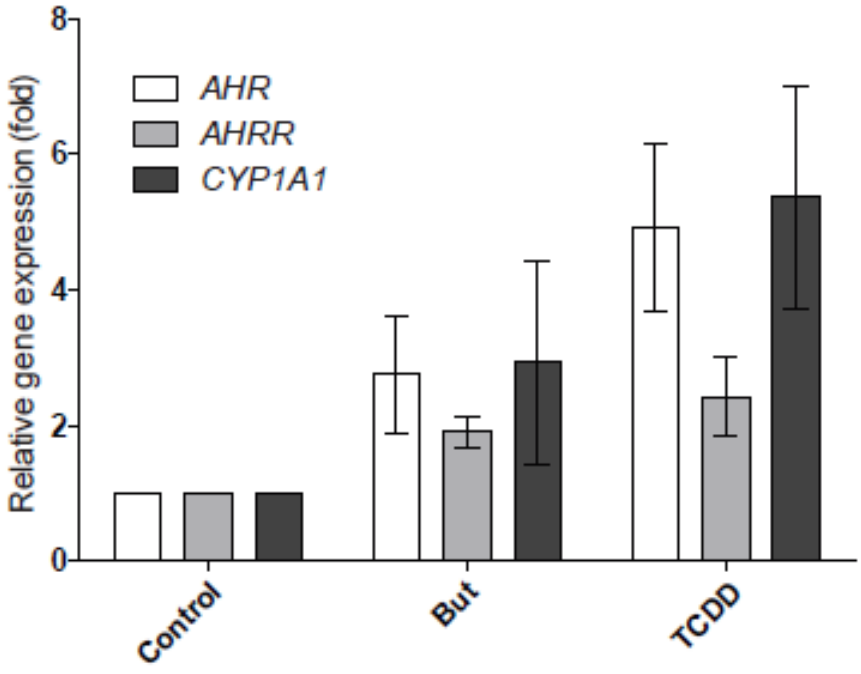
Marinelli *et al.*
Figure 8



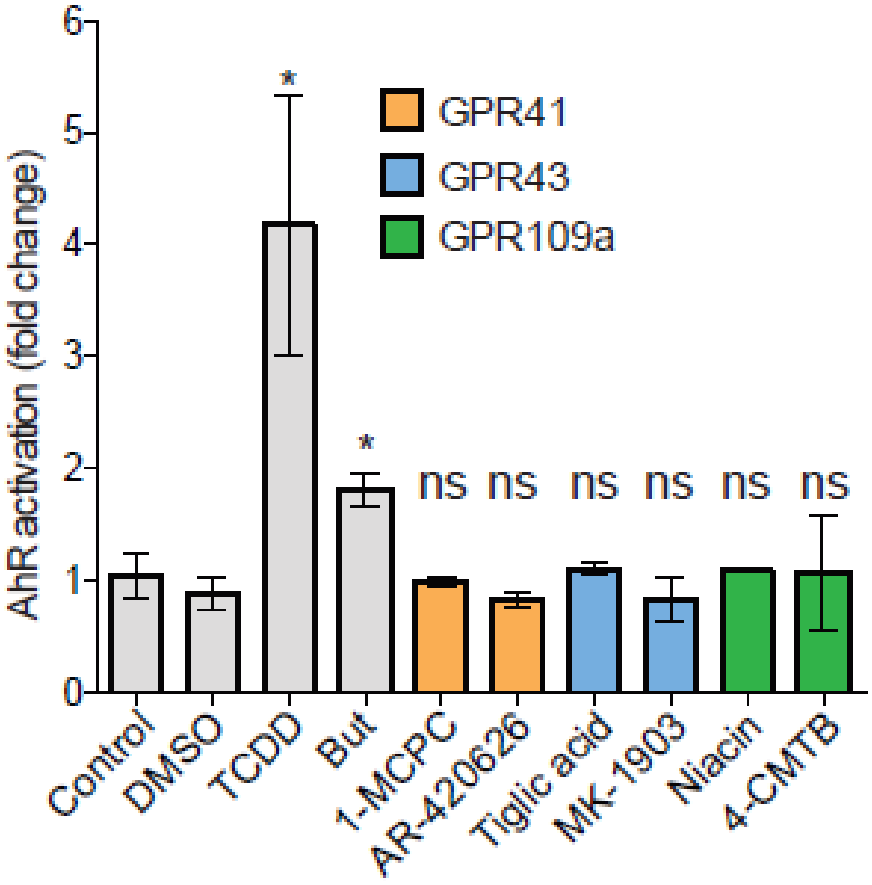
Marinelli *et al.*
Supplementaty Figure S1



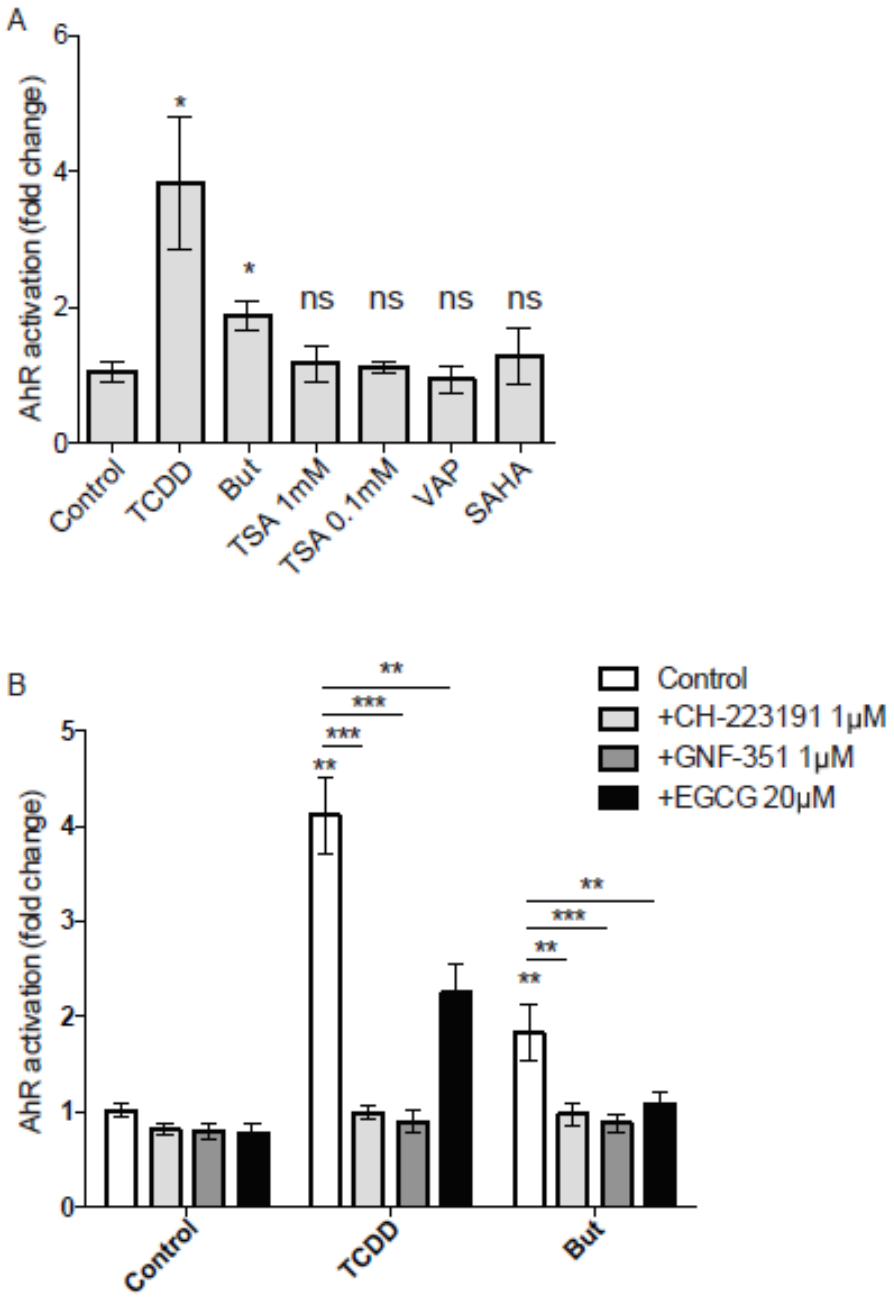
Marinelli *et al.*
Supplementary Figure S2



Marinelli *et al.*
Supplementary Figure S3



Marinelli *et al.*
Supplementary Figure S4



Marinelli *et al.*
Supplementary Figure S5

Marinelli *et al.* Supplementary Table 1

Espece	Aerobic	Phylum	Order	Genus	OD_{600}	Medium	Concentration in mM						AhR activity
							Acetate	Propionate	Isobutyrate	Butyrate	Isovalerate	Valerate	
Adlercreutzia equolifaciens DSM19450, Type strain	No	Actinobacteria	Eggerthellales	Adlercreutzia	0,22	Wilkins-Chalgren Anaerobe Broth	NA	NA	NA	NA	NA	NA	1,38
Adlercreutzia equolifaciens DSM19450, Type strain	No	Actinobacteria	Eggerthellales	Adlercreutzia	0,17	M104 arginin	NA	NA	NA	NA	NA	NA	2,16
Akkermansia muciniphila DSM22959, Type strain	No	Verrucomicrobia	Verrucomicrobiales	Akkermansia	1,07	LYBHI 4 mucin	6,620	4,220	-0,020	-0,070	-0,030	0,000	0,98
Akkermansia muciniphila DSM22959, Type strain	No	Verrucomicrobia	Verrucomicrobiales	Akkermansia	NA	LYBHI 4 mucin	5,090	3,250	-0,080	-0,090	0,000	0,000	1,11
Akkermansia muciniphila DSM22959, Type strain	No	Verrucomicrobia	Verrucomicrobiales	Akkermansia	1,41	LYBHI 4 mucin	NA	NA	NA	NA	NA	NA	1,21
Akkermansia muciniphila DSM22959, Type strain	No	Verrucomicrobia	Verrucomicrobiales	Akkermansia	0,92	LYBHI 4 mucin	NA	NA	NA	NA	NA	NA	1,37
Alistipes shahii DSM19121, Type strain	No	Bacteroidetes	Bacteroidales	Alistipes	1,76	LYBHI 4	NA	NA	NA	NA	NA	NA	1,17
Alistipes shahii DSM19121, Type strain	No	Bacteroidetes	Bacteroidales	Alistipes	0,85	M78	NA	NA	NA	NA	NA	NA	1,17
Alistipes shahii DSM19121, Type strain	No	Bacteroidetes	Bacteroidales	Alistipes	0,37	LYBHI 4 RJ	NA	NA	NA	NA	NA	NA	1,18
Allobaculum stercoricanis DSM13633, Type strain	No	Firmicutes	Erysipelotrichales	Allobaculum	0,26	LYBHI 4 RJ	NA	NA	NA	NA	NA	NA	0,48
Allobaculum stercoricanis DSM13633, Type strain	No	Firmicutes	Erysipelotrichales	Allobaculum	1,28	M104	NA	NA	NA	NA	NA	NA	1,17
Anaerostipes caccae DSM14662, Type strain	No	Firmicutes	Clostridiales	Anaerostipes	0,89	M78 G	-5,130	2,620	0,000	15,930	-0,190	0,270	1,11
Anaerostipes caccae DSM14662, Type strain	No	Firmicutes	Clostridiales	Anaerostipes	NA	LYBHI 4 RJ	NA	NA	NA	NA	NA	NA	1,11
Anaerostipes caccae DSM14662, Type strain	No	Firmicutes	Clostridiales	Anaerostipes	0,74	LYBHI 4 RJ	-10,100	-0,680	0,000	23,190	-0,040	0,110	1,26
Anaerostipes caccae DSM14662, Type strain	No	Firmicutes	Clostridiales	Anaerostipes	NA	LYBHI 4 RJ	NA	NA	NA	NA	NA	NA	1,44
Anaerostipes caccae DSM14662, Type strain	No	Firmicutes	Clostridiales	Anaerostipes	NA	LYBHI 4 RJ	NA	NA	NA	NA	NA	NA	1,48
Anaerostipes caccae DSM14662, Type strain	No	Firmicutes	Clostridiales	Anaerostipes	NA	LYBHI 4 RJ	NA	NA	NA	NA	NA	NA	1,48
Anaerostipes caccae DSM14662, Type strain	No	Firmicutes	Clostridiales	Anaerostipes	1,58	M104	-4,330	0,000	0,000	14,460	0,000	0,000	1,64
Anaerostipes caccae DSM14662, Type strain	No	Firmicutes	Clostridiales	Anaerostipes	NA	M104	NA	NA	NA	NA	NA	NA	2,28
Anaerostipes hadrus DSM3319, Type strain	No	Firmicutes	Clostridiales	Anaerostipes	NA	M78	-8,830	-0,085	-0,035	4,215	-0,043	-0,025	1,07
Anaerostipes hadrus DSM3319, Type strain	No	Firmicutes	Clostridiales	Anaerostipes	0,60	M78	-8,080	-0,035	-0,015	4,015	0,068	-0,025	1,38
Anaerostipes hadrus DSM3319, Type strain	No	Firmicutes	Clostridiales	Anaerostipes	1,94	M110	-2,050	-0,115	-0,045	17,525	0,248	-0,025	1,70
Anaerotruncus colihominis DSM17241, Type strain	No	Firmicutes	Clostridiales	Anaerotruncus	NA	LYBHI 4 RJ	12,240	1,410	0,190	9,060	0,350	4,094	1,58
Anaerotruncus colihominis DSM17241, Type strain	No	Firmicutes	Clostridiales	Anaerotruncus	NA	LYBHI 4 RJ	14,840	1,750	0,180	8,600	0,370	4,084	1,90
Anaerotruncus colihominis DSM17241, Type strain	No	Firmicutes	Clostridiales	Anaerotruncus	1,05	LYBHI 4 RJ	6,940	0,000	0,000	8,620	0,080	3,894	2,31
Atopobium parvulum DSM20469, Type strain	No	Actinobacteria	Coriobacteriales	Atopobium	0,50	Wilkins-Chalgren Anaerobe Broth	NA	NA	NA	NA	NA	NA	0,63
Bacteroides caccae DSM19024, Type strain	No	Bacteroidetes	Bacteroidales	Bacteroides	0,49	LYHBHI 1	9,690	9,990	0,090	-0,220	1,850	0,000	1,34
Bacteroides cellulosilyticus DSM14838, Type strain	No	Bacteroidetes	Bacteroidales	Bacteroides	0,26	M78	NA	NA	NA	NA	NA	NA	1,30
Bacteroides cellulosilyticus DSM14838, Type strain	No	Bacteroidetes	Bacteroidales	Bacteroides	1,48	LYBHI 4 RJ	NA	NA	NA	NA	NA	NA	1,37
Bacteroides fragilis DSM2151, Type strain	No	Bacteroidetes	Bacteroidales	Bacteroides	0,50	LYHBHI 1	10,500	13,830	0,320	-0,180	3,390	0,000	1,30
Bacteroides ovatus DSM1896, Type strain	No	Bacteroidetes	Bacteroidales	Bacteroides	0,68	M104	9,690	0,030	-1,050	-0,240	0,070	0,000	1,17
Bacteroides spp. D8	No	Bacteroidetes	Bacteroidales	Bacteroides	0,74	M104	10,570	0,090	-1,830	-0,240	0,030	0,000	0,99
Bacteroides stercoris DSM19555, Type strain	No	Bacteroidetes	Bacteroidales	Bacteroides	0,87	M78	NA	NA	NA	NA	NA	NA	0,80
Bacteroides thetaiotaomicron DSM2079, Type strain	No	Bacteroidetes	Bacteroidales	Bacteroides	0,77	M104	9,920	0,050	-1,790	-0,240	0,060	0,000	0,97
Bacteroides thetaiotaomicron DSM2079, Type strain	No	Bacteroidetes	Bacteroidales	Bacteroides	NA	M17	NA	NA	NA	NA	NA	NA	1,05
Bacteroides thetaiotaomicron DSM2079, Type strain	No	Bacteroidetes	Bacteroidales	Bacteroides	0,30	M110	NA	NA	NA	NA	NA	NA	1,24
Bacteroides uniformis DSM6597, Type strain	No	Bacteroidetes	Bacteroidales	Bacteroides	1,09	M104	9,770	46,690	-1,330	-0,190	0,280	0,000	1,12
Bacteroides uniformis DSM6597, Type strain	No	Bacteroidetes	Bacteroidales	Bacteroides	0,70	LYBHI 4 RJ	NA	NA	NA	NA	NA	NA	1,15
Bacteroides uniformis DSM6597, Type strain	No	Bacteroidetes	Bacteroidales	Bacteroides	NA	LYBHI 4	NA	NA	NA	NA	NA	NA	1,24
Bacteroides vulgatus DSM1447, Type strain	No	Bacteroidetes	Bacteroidales	Bacteroides	0,44	LYHBHI 1	13,730	10,070	0,260	-0,230	2,460	0,000	1,15
Bacteroides xylanisolvens DSM18836, Type strain	No	Bacteroidetes	Bacteroidales	Bacteroides	0,94	LYBHI 4 RJ	NA	NA	NA	NA	NA	NA	1,01
Bacteroides xylanisolvens DSM18836, Type strain	No	Bacteroidetes	Bacteroidales	Bacteroides	1,06	LYBHI 4	NA	NA	NA	NA	NA	NA	1,18
Bifidobacterium adolescentis DSM20083, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	0,88	M58	73,420	0,000	-1,030	0,020	0,000	0,000	1,32
Bifidobacterium adolescentis DSM20086	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	2,00	M58	NA	NA	NA	NA	NA	NA	1,64
Bifidobacterium adolescentis DSM20086	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	1,48	LYBHI 4	NA	NA	NA	NA	NA	NA	1,95
Bifidobacterium angulatum DSM20098, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	1,07	M58	69,210	0,000	-1,100	-0,010	0,000	0,000	1,19
Bifidobacterium animalis subsp. animalis DSM20104, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	0,68	M58	49,850	0,000	-0,360	0,060	0,000	0,000	1,56
Bifidobacterium animalis subsp. animalis DSM20104, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	NA	M58	NA	NA	NA	NA	NA	NA	1,64
Bifidobacterium bifidum DSM20082	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	0,79	M58	69,390	0,000	-0,900	-0,010	0,000	0,000	1,31
Bifidobacterium bifidum DSM20082	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	NA	M58	NA	NA	NA	NA	NA	NA	1,42
Bifidobacterium breve ATCC15701	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	0,31	M104	15,610	0,000	-1,400	-0,180	0,000	0,000	1,71
Bifidobacterium breve ATCC15701	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	NA	M58	NA	NA	NA	NA	NA	NA	1,89
Bifidobacterium breve DSM20091	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	0,52	M104	1,130	12,520	-1,020	-0,190	0,240	0,000	1,04
Bifidobacterium catenulatum DSM16992, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	1,07	M58	62,140	0,000	-0,810	-0,010	0,000	0,000	1,33
Bifidobacterium choerinum DSM20434, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	NA	M58	NA	NA	NA	NA	NA	NA	1,68
Bifidobacterium choerinum DSM20434, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	0,97	M58	69,050	0,000	-1,040	0,000	0,000	0,000	1,84
Bifidobacterium dentium DSM20084	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	0,88	M104	48,720	0,000	-1,840	-0,220	0,000	0,000	1,18
Bifidobacterium dentium DSM20436, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	0,80	M104	42,870	0,000	-1,660	-0,180	0,000	0,000	1,78
Bifidobacterium gallicum DSM20093, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	0,58	M58	51,140	0,000	-0,570	0,000	0,000	0,000	1,09
Bifidobacterium longum subsp. infantis DSM20088, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	0,76	M58	49,770	0,000	-1,010	-0,010	0,000	0,000	1,37
Bifidobacterium longum subsp. infantis DSM20088, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	1,32	M58	NA	NA	NA	NA	NA	NA	2,17
Bifidobacterium longum subsp. longum DSM20219, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	0,90	M58	65,040	0,000	-0,840	-0,020	0,000	0,000	1,48

Bifidobacterium longum subsp. longum DSM 20219, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	NA	M58	NA	NA	NA	NA	NA	2,27
Bifidobacterium pseudocatenulatum DSM 20438, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	1,27	M58	71,020	0,000	-1,070	-0,020	0,000	1,26
Bifidobacterium mtuiantium DSM 6489, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	1,08	M58	61,090	0,000	-1,150	-0,010	0,000	1,14
Blautia coccoides DSM 935, Type strain	No	Firmicutes	Clostridiales	Blautia	1,94	LYBH11	32,580	-0,130	-0,060	-0,060	-0,080	0,88
Blautia coccoides DSM 935, Type strain	No	Firmicutes	Clostridiales	Blautia	1,69	LYBH11	39,370	-0,140	-0,060	-0,080	-0,090	0,88
Blautia coccoides DSM 935, Type strain	No	Firmicutes	Clostridiales	Blautia	1,88	LYBH11 mucin	40,220	-0,130	0,020	-0,050	-0,080	0,92
Blautia coccoides DSM 935, Type strain	No	Firmicutes	Clostridiales	Blautia	NA	LYBH11	NA	NA	NA	NA	NA	0,93
Blautia coccoides DSM 935, Type strain	No	Firmicutes	Clostridiales	Blautia	0,79	M78	21,000	-0,415	-0,075	-0,415	-0,053	3,842
Blautia coccoides DSM 935, Type strain	No	Firmicutes	Clostridiales	Blautia	1,25	M78	20,320	-0,395	-0,075	-0,415	-0,083	3,842
Blautia coccoides DSM 935, Type strain	No	Firmicutes	Clostridiales	Blautia	1,78	M110	34,690	-0,365	-0,075	-0,415	0,138	3,842
Blautia coccoides DSM 935, Type strain	No	Firmicutes	Clostridiales	Blautia	0,79	M78	37,140	-0,375	-0,015	-0,225	-0,133	3,842
Blautia coccoides DSM 935, Type strain	No	Firmicutes	Clostridiales	Blautia	1,49	M78	21,720	-0,365	-0,075	-0,275	-0,073	3,842
Blautia coccoides DSM 935, Type strain	No	Firmicutes	Clostridiales	Blautia	1,91	LYBH14 RJ	46,250	0,500	0,020	0,200	0,020	0,010
Blautia gnavus, B07	No	Firmicutes	Clostridiales	Blautia	0,89	LYBH11	9,600	-0,070	0,010	-0,030	-0,060	0,000
Blautia gnavus, B07	No	Firmicutes	Clostridiales	Blautia	1,21	LYBH14	26,570	3,450	0,370	1,870	0,250	0,330
Blautia gnavus, B07	No	Firmicutes	Clostridiales	Blautia	1,45	LYBH14	28,060	3,240	0,360	1,820	0,650	0,350
Blautia gnavus, B07	No	Firmicutes	Clostridiales	Blautia	1,45	LYBH14	NA	NA	NA	NA	NA	NA
Blautia gnavus, B09	No	Firmicutes	Clostridiales	Blautia	NA	LYBH11	NA	NA	NA	NA	NA	NA
Blautia gnavus, B09	No	Firmicutes	Clostridiales	Blautia	1,44	LYBH14	24,540	3,440	0,370	1,850	0,690	0,330
Blautia gnavus, B94	No	Firmicutes	Clostridiales	Blautia	NA	LYBH14	29,000	3,900	0,420	2,090	0,330	0,370
Blautia gnavus, B94	No	Firmicutes	Clostridiales	Blautia	1,56	LYBH14	NA	NA	NA	NA	NA	NA
Blautia gnavus, B94	No	Firmicutes	Clostridiales	Blautia	NA	LYBH14	NA	NA	NA	NA	NA	NA
Blautia gnavus, B94	No	Firmicutes	Clostridiales	Blautia	1,12	LYBH11	12,380	0,160	0,030	0,100	0,080	0,000
Blautia gnavus, B94	No	Firmicutes	Clostridiales	Blautia	1,56	LYBH14	26,170	3,350	0,360	1,810	0,680	0,320
Blautia gnavus, FRE1	No	Firmicutes	Clostridiales	Blautia	1,61	LYBH14	11,270	0,000	0,000	-0,010	0,000	0,000
Blautia gnavus, FRE1	No	Firmicutes	Clostridiales	Blautia	1,63	LYBH14	11,830	0,020	0,000	0,020	0,040	0,000
Blautia gnavus, FRE1	No	Firmicutes	Clostridiales	Blautia	1,79	LYBH14 RJ	14,570	0,330	0,020	-0,050	0,010	-0,120
Blautia gnavus, FRE1	No	Firmicutes	Clostridiales	Blautia	1,91	LYBH14	13,960	0,000	-0,010	0,010	-0,400	0,000
Blautia gnavus, FRE1	No	Firmicutes	Clostridiales	Blautia	1,63	LYBH14	11,830	0,020	0,000	0,020	0,040	0,000
Blautia gnavus, V62	No	Firmicutes	Clostridiales	Blautia	1,62	LYBH14	10,540	29,830	0,370	1,720	0,340	0,270
Blautia gnavus, V62	No	Firmicutes	Clostridiales	Blautia	NA	LYBH14	6,600	0,080	-0,020	-0,070	0,400	0,000
Blautia gnavus, V62	No	Firmicutes	Clostridiales	Blautia	1,86	LYBH14	15,450	37,140	0,350	1,700	0,830	0,320
Blautia gnavus, V62	No	Firmicutes	Clostridiales	Blautia	1,86	LYBH14	15,450	37,140	0,350	1,700	0,830	0,320
Blautia gnavus, V62	No	Firmicutes	Clostridiales	Blautia	1,62	LYBH14	10,410	29,210	0,340	1,630	0,310	0,350
Blautia gnavus, V96	No	Firmicutes	Clostridiales	Blautia	NA	LYBH14	3,170	0,720	-0,010	0,020	0,060	0,000
Blautia gnavus, V96	No	Firmicutes	Clostridiales	Blautia	NA	LYBH11	3,150	3,000	-0,010	-0,060	0,040	0,000
Blautia gnavus, V96	No	Firmicutes	Clostridiales	Blautia	1,55	LYBH14	15,460	34,580	0,350	1,680	0,440	0,290
Blautia gnavus, V96	No	Firmicutes	Clostridiales	Blautia	1,96	LYBH14	19,570	41,630	0,380	1,750	0,970	0,290
Blautia gnavus, V96	No	Firmicutes	Clostridiales	Blautia	1,55	LYBH14	15,460	34,580	0,350	1,680	0,440	0,290
Blautia gnavus, V96	No	Firmicutes	Clostridiales	Blautia	1,96	LYBH14	19,570	41,630	0,380	1,750	0,970	0,290
Blautia hansenii DSM 20583, Type strain	No	Firmicutes	Clostridiales	Blautia	1,54	M104	10,610	-0,170	-0,020	-0,180	-0,030	0,000
Blautia hansenii DSM 20583, Type strain	No	Firmicutes	Clostridiales	Blautia	1,54	M104	79,580	19,820	1,810	10,060	3,960	1,670
Blautia hansenii DSM 20583, Type strain	No	Firmicutes	Clostridiales	Blautia	NA	M78	NA	NA	NA	NA	NA	1,47
Blautia hansenii DSM 20583, Type strain	No	Firmicutes	Clostridiales	Blautia	1,59	M104	3,770	0,000	0,000	0,000	0,000	2,33
Blautia hansenii DSM 20583, Type strain	No	Firmicutes	Clostridiales	Blautia	1,30	LYBH14	9,190	0,000	0,000	0,020	0,000	1,22
Blautia hydrogenotrophica DSM 10507, Type strain	No	Firmicutes	Clostridiales	Blautia	NA	LYBH14 RJ	16,140	0,780	0,520	0,320	1,290	1,03
Blautia hydrogenotrophica DSM 10507, Type strain	No	Firmicutes	Clostridiales	Blautia	0,60	M78	3,150	0,145	0,545	-0,005	1,358	-0,025
Blautia hydrogenotrophica DSM 10507, Type strain	No	Firmicutes	Clostridiales	Blautia	1,50	LYBH14 RJ	30,350	-0,080	0,260	-0,010	0,450	0,030
Blautia hydrogenotrophica DSM 10507, Type strain	No	Firmicutes	Clostridiales	Blautia	1,39	LYBH14	21,130	0,050	1,080	0,030	1,990	0,000
Blautia hydrogenotrophica DSM 10507, Type strain	No	Firmicutes	Clostridiales	Blautia	NA	LYBH14	18,060	0,000	0,540	0,000	1,480	0,000
Blautia hydrogenotrophica DSM 10507, Type strain	No	Firmicutes	Clostridiales	Blautia	1,58	M104	14,660	0,000	0,000	0,000	0,000	1,41
Blautia hydrogenotrophica DSM 10507, Type strain	No	Firmicutes	Clostridiales	Blautia	NA	LYBH14 RJ	30,810	-0,120	4,570	-0,120	9,290	-0,120
Blautia obeum DSM 25238, Type strain	No	Firmicutes	Clostridiales	Blautia	1,47	LYBH14	24,750	-0,020	-0,010	0,030	0,060	0,000
Blautia obeum DSM 25238, Type strain	No	Firmicutes	Clostridiales	Blautia	1,79	LYBH14	27,780	0,130	0,020	0,150	0,040	0,000
Blautia obeum DSM 25238, Type strain	No	Firmicutes	Clostridiales	Blautia	1,95	LYBH14	28,490	0,380	0,070	0,280	0,160	0,000
Blautia obeum DSM 25238, Type strain	No	Firmicutes	Clostridiales	Blautia	2,00	LYBH14 RJ	28,620	0,380	0,060	0,160	0,110	0,030
Blautia producta DSM 2950, Type strain	No	Firmicutes	Clostridiales	Blautia	1,03	M104	0,190	-0,150	-0,080	-0,180	0,020	0,000
Blautia producta DSM 2950, Type strain	No	Firmicutes	Clostridiales	Blautia	1,38	LYBH14 RJ	3,020	-0,060	-0,020	-0,230	-0,040	-0,120
Blautia producta DSM 2950, Type strain	No	Firmicutes	Clostridiales	Blautia	1,94	M104	2,580	-0,160	-0,070	-0,180	0,040	0,000
Blautia producta DSM 2950, Type strain	No	Firmicutes	Clostridiales	Blautia	1,31	LYBH14	3,210	-0,180	-0,080	-0,110	-0,030	0,000
Blautia producta DSM 2950, Type strain	No	Firmicutes	Clostridiales	Blautia	0,57	M78	-4,610	-0,375	-0,075	-0,415	-0,023	-0,025
Blautia producta DSM 2950, Type strain	No	Firmicutes	Clostridiales	Blautia	NA	M104	NA	NA	NA	NA	NA	1,28
Blautia wexlerae DSM 19850, Type strain	No	Firmicutes	Clostridiales	Blautia	0,27	M104	15,880	3,520	0,370	1,810	0,690	0,350
Blautia wexlerae DSM 19850, Type strain	No	Firmicutes	Clostridiales	Blautia	0,39	M78	-4,420	0,185	0,025	0,015	0,128	-0,025
Blautia wexlerae DSM 19850, Type strain	No	Firmicutes	Clostridiales	Blautia	0,41	M78	11,210	3,535	0,395	1,725	0,688	0,375
Butyrivibrio fibrisolvens DSM 3071, Type strain	No	Firmicutes	Clostridiales	Butyrivibrio	0,66	LYBH12	1,640	0,000	-0,070	-0,050	0,020	0,000

Butyrivibrio fibrisolvens DSM3071, Type strain	No	Firmicutes	Clostridiales	Butyrivibrio	0,79	LYBHI 4 RJ	4,110	0,390	0,000	8,300	0,000	0,090	1,21
Butyrivibrio fibrisolvens DSM3071, Type strain	No	Firmicutes	Clostridiales	Butyrivibrio	0,40	LYBHI 4 RJ	104,860	30,280	3,200	15,290	6,000	2,840	1,24
Clostridium paraputrificum, G12PR - X73445	No	Firmicutes	Clostridiales	Clostridium	0,20	LYHBHI 1	7,750	0,000	0,150	3,710	0,000	0,000	1,37
Clostridium sardiniense, C3	No	Firmicutes	Clostridiales	Clostridium	0,31	M78	2,920	0,160	0,020	7,720	0,000	0,000	2,05
Clostridium sardiniense, C7	No	Firmicutes	Clostridiales	Clostridium	0,39	M78	5,540	0,430	0,000	15,230	0,000	0,000	2,05
Clostridium sardiniense, C9	No	Firmicutes	Clostridiales	Clostridium	0,29	M78	7,060	0,440	0,000	7,920	0,000	0,000	2,04
Clostridium sardiniense, D12	No	Firmicutes	Clostridiales	Clostridium	0,40	M78	2,530	0,050	0,000	6,150	0,000	0,000	1,90
Clostridium sardiniense, D7	No	Firmicutes	Clostridiales	Clostridium	0,57	M78	6,530	0,000	-0,100	-0,200	0,000	0,000	2,31
Clostridium sardiniense, I11	No	Firmicutes	Clostridiales	Clostridium	0,34	M78	2,170	0,000	7,970	13,220	0,000	0,000	1,78
Collinsella aerofaciens DSM3979, Type strain	No	Actinobacteria	Coriobacteriales	Collinsella	0,19	LYHBHI 1	3,960	0,000	-0,150	-0,170	0,000	0,000	1,07
Collinsella aerofaciens DSM3979, Type strain	No	Actinobacteria	Coriobacteriales	Collinsella	NA	Wilkins-Chalgren Anaerobe Broth	9,570	0,000	0,110	0,010	0,030	0,000	1,10
Coprococcus eutactus ATCC27759, Type strain	No	Firmicutes	Clostridiales	Coprococcus	0,41	M78	0,960	1,355	-0,015	0,135	0,178	-0,025	1,27
Coprococcus eutactus ATCC27759, Type strain	No	Firmicutes	Clostridiales	Coprococcus	0,97	M110	-4,490	12,375	-0,035	-0,215	0,488	-0,025	1,33
Dorea formicigenerans DSM3992, Type strain	No	Firmicutes	Clostridiales	Dorea	0,56	LYBHI 1	2,440	0,650	-0,080	2,390	0,000	0,000	1,36
Dorea formicigenerans DSM3992, Type strain	No	Firmicutes	Clostridiales	Dorea	0,86	M110	-1,820	0,555	-0,045	6,525	0,138	-0,025	1,63
Dorea formicigenerans DSM3992, Type strain	No	Firmicutes	Clostridiales	Dorea	1,45	M78	7,260	19,225	-0,075	4,645	0,548	-0,025	1,71
Dorea longicatena DSM13814, Type strain	No	Firmicutes	Clostridiales	Dorea	1,58	LYBHI 4	11,210	0,000	0,000	0,020	0,000	0,000	0,81
Dorea longicatena DSM13814, Type strain	No	Firmicutes	Clostridiales	Dorea	1,37	LYBHI 4 RJ	12,480	-0,120	0,000	-0,050	-0,050	0,050	0,91
Dorea longicatena DSM13814, Type strain	No	Firmicutes	Clostridiales	Dorea	2,00	M104	12,910	0,000	0,000	0,000	0,000	0,000	1,93
Edwardsiella tarda ATCC23685	S	Proteobacteria	Gammaproteobacteria	Edwardsiella	NA	M3	NA	NA	NA	NA	NA	NA	1,06
Edwardsiella tarda ATCC23685	S	Proteobacteria	Gammaproteobacteria	Edwardsiella	NA	M104	NA	NA	NA	NA	NA	NA	1,27
Edwardsiella tarda ATCC23685	S	Proteobacteria	Gammaproteobacteria	Edwardsiella	NA	M3	NA	NA	NA	NA	NA	NA	1,36
Edwardsiella tarda DSM30052, Type strain	S	Proteobacteria	Gammaproteobacteria	Edwardsiella	NA	LYBHI 4	NA	NA	NA	NA	NA	NA	1,04
Edwardsiella tarda DSM30052, Type strain	S	Proteobacteria	Gammaproteobacteria	Edwardsiella	NA	LYBHI 4 RJ	NA	NA	NA	NA	NA	NA	1,12
Edwardsiella tarda DSM30052, Type strain	S	Proteobacteria	Gammaproteobacteria	Edwardsiella	1,34	LYBHI 1	NA	NA	NA	NA	NA	NA	1,21
Edwardsiella tarda DSM30052, Type strain	S	Proteobacteria	Gammaproteobacteria	Edwardsiella	0,73	LB	NA	NA	NA	NA	NA	NA	1,25
Edwardsiella tarda DSM30052, Type strain	S	Proteobacteria	Gammaproteobacteria	Edwardsiella	0,44	M1	NA	NA	NA	NA	NA	NA	1,27
Edwardsiella tarda DSM30052, Type strain	S	Proteobacteria	Gammaproteobacteria	Edwardsiella	0,85	M1	NA	NA	NA	NA	NA	NA	1,28
Edwardsiella tarda DSM30052, Type strain	S	Proteobacteria	Gammaproteobacteria	Edwardsiella	0,29	M1	NA	NA	NA	NA	NA	NA	1,29
Edwardsiella tarda DSM30052, Type strain	S	Proteobacteria	Gammaproteobacteria	Edwardsiella	0,36	M1	NA	NA	NA	NA	NA	NA	1,31
Edwardsiella tarda DSM30052, Type strain	S	Proteobacteria	Gammaproteobacteria	Edwardsiella	0,82	LB	NA	NA	NA	NA	NA	NA	1,32
Edwardsiella tarda DSM30052, Type strain	S	Proteobacteria	Gammaproteobacteria	Edwardsiella	3,16	LB	NA	NA	NA	NA	NA	NA	1,35
Edwardsiella tarda DSM30052, Type strain	S	Proteobacteria	Gammaproteobacteria	Edwardsiella	1,46	M1	NA	NA	NA	NA	NA	NA	1,62
Edwardsiella tarda DSM30052, Type strain	S	Proteobacteria	Gammaproteobacteria	Edwardsiella	NA	M104	NA	NA	NA	NA	NA	NA	1,74
Eggerthella lenta DSM2243, Type strain	No	Actinobacteria	Eggerthellales	Eggerthella	NA	Wilkins-Chalgren Anaerobe Broth	1,710	0,000	0,000	0,090	0,120	0,000	0,72
Enterococcus faecium, FAIR E345	Yes	Firmicutes	Lactobacillales	Enterococcus	NA	BHI aerobic	2,500	0,000	0,010	0,010	-0,010	0,000	1,05
Erysipelatoclostridium ramosum DSM1402, Type strain	No	Firmicutes	Clostridiales	Erysipelatoclostridium	1,43	LYBHI 1	14,290	0,230	0,040	0,14	-0,02	0,00	1,02
Erysipelatoclostridium ramosum DSM1402, Type strain	No	Firmicutes	Clostridiales	Erysipelatoclostridium	1,78	LYBHI 4	14,580	0,150	0,010	0,06	0,06	0,00	1,11
Erysipelatoclostridium ramosum DSM1402, Type strain	No	Firmicutes	Clostridiales	Erysipelatoclostridium	1,57	LYBHI 4 RJ	16,790	0,280	0,000	-0,10	-0,01	-0,12	1,20
Erysipelatoclostridium ramosum DSM1402, Type strain	No	Firmicutes	Clostridiales	Erysipelatoclostridium	1,19	LYBHI 4 RJ	23,020	0,050	0,000	10,44	0,02	0,01	1,36
Erysipelatoclostridium saccharogumia DSM17460, Type strain	No	Firmicutes	Clostridiales	Erysipelatoclostridium	0,58	M104 G	6,840	0,000	0,000	0,00	0,00	0,00	1,20
Erysipelatoclostridium saccharogumia DSM17460, Type strain	No	Firmicutes	Clostridiales	Erysipelatoclostridium	1,22	M104	3,830	0,000	0,000	0,00	0,00	0,00	1,56
Erysipelotrichaceae spiroforme DSM1552, Type strain	No	Firmicutes	Clostridiales	Erysipelatoclostridium	0,77	M110	8,400	0,000	0,000	0,07	-0,33	0,00	1,13
Erysipelotrichaceae spiroforme DSM1552, Type strain	No	Firmicutes	Clostridiales	Erysipelatoclostridium	0,97	LYBHI 4 RJ	73,610	5,630	0,440	3,83	0,51	0,33	1,15
Erysipelotrichaceae spiroforme DSM1552, Type strain	No	Firmicutes	Clostridiales	Erysipelatoclostridium	0,99	LYBHI 4	8,420	0,080	0,000	0,17	0,04	0,00	1,34
Escherichia coli, LF82	Yes	Proteobacteria	Gammaproteobacteria	Escherichia	NA	LB	NA	NA	NA	NA	NA	NA	2,45
Escherichia coli, LF82	Yes	Proteobacteria	Gammaproteobacteria	Escherichia	NA	LB	NA	NA	NA	NA	NA	NA	2,48
Escherichia coli, MG1655	Yes	Proteobacteria	Gammaproteobacteria	Escherichia	NA	LB	NA	NA	NA	NA	NA	NA	2,15
Escherichia coli, S123	Yes	Proteobacteria	Gammaproteobacteria	Escherichia	NA	LB	NA	NA	NA	NA	NA	NA	2,28
Escherichia coli, S123	Yes	Proteobacteria	Gammaproteobacteria	Escherichia	NA	LB	NA	NA	NA	NA	NA	NA	2,68
Eubacterium coprostanoligenes ATCC51222, Type strain	No	Firmicutes	Clostridiales	Eubacterium	NA	LYBHI 4 CH/PL	NA	NA	NA	NA	NA	NA	0,89
Eubacterium coprostanoligenes ATCC51222, Type strain	No	Firmicutes	Clostridiales	Eubacterium	NA	BCM	NA	NA	NA	NA	NA	NA	1,06
Eubacterium eligens DSM3376, Type strain	No	Firmicutes	Clostridiales	Eubacterium	0,49	M436	-0,510	0,000	0,650	-1,00	-0,91	0,00	1,06
Eubacterium eligens DSM3376, Type strain	No	Firmicutes	Clostridiales	Eubacterium	1,03	LYBHI 4 RJ	-3,710	-1,320	1,000	-0,86	1,65	0,00	1,06
Eubacterium eligens DSM3376, Type strain	No	Firmicutes	Clostridiales	Eubacterium	0,82	LYBHI 4	NA	NA	NA	NA	NA	NA	1,08
Eubacterium eligens DSM3376, Type strain	No	Firmicutes	Clostridiales	Eubacterium	1,30	M104	3,750	0,000	0,000	0,00	0,01	0,00	1,20
Eubacterium hallii DSM3353, Type strain	No	Firmicutes	Clostridiales	Eubacterium	0,38	M78	-5,960	-0,205	-0,045	-0,245	-0,013	-0,025	0,98
Eubacterium limosum DSM20543, Type strain	No	Firmicutes	Clostridiales	Eubacterium	NA	LYBHI 3	13,870	0,040	0,420	6,200	0,410	0,110	1,65
Eubacterium limosum DSM20543, Type strain	No	Firmicutes	Clostridiales	Eubacterium	2,22	M104	19,200	-0,050	-0,040	13,460	0,020	0,110	2,44
Eubacterium limosum DSM20543, Type strain	No	Firmicutes	Clostridiales	Eubacterium	NA	M104	23,460	0,050	-0,020	12,310	0,010	0,000	3,20
Eubacterium rectale ATCC33656 - CIP105953	No	Firmicutes	Clostridiales	Eubacterium	0,23	LYHBHI 1	0,100	4,640	-0,150	-0,160	0,000	0,000	1,11
Eubacterium rectale ATCC33656 - CIP105953	No	Firmicutes	Clostridiales	Eubacterium	0,23	LYHBHI 4	NA	NA	NA	NA	NA	NA	1,41
Eubacterium ventriosum DSM3988, Type strain	No	Firmicutes	Clostridiales	Eubacterium	0,27	M110	-0,690	0,515	-0,005	0,125	0,168	-0,025	1,16
Eubacterium ventriosum DSM3988, Type strain	No	Firmicutes	Clostridiales	Eubacterium	0,96	LYBHI 4 RJ	-1,280	4,470	0,000	-0,100	-0,020	0,020	1,19
Eubacterium ventriosum DSM3988, Type strain	No	Firmicutes	Clostridiales	Eubacterium	NA	M110 K1	8,010	29,720	0,000	0,000	0,820	0,000	1,25
Faecalibacterium prausnitzii ATCC27768, Type strain	No	Firmicutes	Clostridiales	Faecalibacterium	0,42	LYBHI 4 RJ	NA	NA	NA	NA	NA	NA	1,11

Faecalibacterium prausnitzii ATCC27768, Type strain	No	Firmicutes	Clostridiales	Faecalibacterium	0,64	M110	NA	NA	NA	NA	NA	NA	NA	1,29
Faecalibacterium prausnitzii ATCC27768, Type strain	No	Firmicutes	Clostridiales	Faecalibacterium	0,61	Wilkins-Chalgren Anaerobe Broth	NA	NA	NA	NA	NA	NA	NA	1,45
Faecalibacterium prausnitzii ATCC27768, Type strain	No	Firmicutes	Clostridiales	Faecalibacterium	0,85	LYBHI 4	NA	NA	NA	NA	NA	NA	NA	1,51
Faecalibacterium prausnitzii DSM17677, Strain A2-165	No	Firmicutes	Clostridiales	Faecalibacterium	1,18	LYBHI 2	-1,210	-0,080	-0,020	7,780	-0,010	0,000		2,05
Faecalibacterium prausnitzii DSM17677, Strain A2-165	No	Firmicutes	Clostridiales	Faecalibacterium	1,70	LYBHI 4	18,640	-0,030	-0,020	-0,030	-0,460	0,000		0,98
Faecalibacterium prausnitzii DSM17677, Strain A2-165	No	Firmicutes	Clostridiales	Faecalibacterium	1,66	LYBHI 4	-2,740	0,000	0,000	7,480	0,010	0,000		1,19
Faecalibacterium prausnitzii DSM17677, Strain A2-165	No	Firmicutes	Clostridiales	Faecalibacterium	NA	LYBHI 4	NA	NA	NA	NA	NA	NA		1,30
Faecalibacterium prausnitzii DSM17677, Strain A2-165	No	Firmicutes	Clostridiales	Faecalibacterium	NA	LYBHI 4 RU	NA	NA	NA	NA	NA	NA		1,34
Faecalibacterium prausnitzii DSM17677, Strain A2-165	No	Firmicutes	Clostridiales	Faecalibacterium	NA	LYBHI 4 RU	-9,900	-1,320	0,000	6,110	-0,110	0,000		1,36
Faecalibacterium prausnitzii DSM17677, Strain A2-165	No	Firmicutes	Clostridiales	Faecalibacterium	0,59	LYBHI 4	-5,520	0,000	0,000	8,630	0,020	0,000		1,56
Faecalibacterium prausnitzii DSM17677, Strain A2-165	No	Firmicutes	Clostridiales	Faecalibacterium	0,36	LYBHI 4	-0,490	1,340	0,000	15,310	0,140	0,080		1,60
Faecalibacterium prausnitzii DSM17677, Strain A2-165	No	Firmicutes	Clostridiales	Faecalibacterium	1,60	LYBHI 4	18,130	0,110	-0,020	-0,050	0,350	0,000		1,60
Faecalibacterium prausnitzii DSM17677, Strain A2-165	No	Firmicutes	Clostridiales	Faecalibacterium	NA	LYBHI 4	NA	NA	NA	NA	NA	NA		1,62
Faecalibacterium prausnitzii DSM17677, Strain A2-165	No	Firmicutes	Clostridiales	Faecalibacterium	0,90	LYBHI 4	-0,830	-0,020	-0,020	8,020	-0,060	0,000		1,84
Faecalibacterium prausnitzii, L2-6	No	Firmicutes	Clostridiales	Faecalibacterium	NA	LYBHI 4	NA	NA	NA	NA	NA	NA		1,01
Faecalibacterium prausnitzii, L2-6	No	Firmicutes	Clostridiales	Faecalibacterium	0,91	LYBHI 4	8,350	0,050	0,000	0,000	0,010	0,000		0,89
Faecalibacterium prausnitzii, L2-6	No	Firmicutes	Clostridiales	Faecalibacterium	NA	LYBHI 4	5,220	0,000	0,000	6,880	0,000	0,000		1,02
Faecalitalea contorta DSM3982, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,92	M110	-0,460	0,425	-0,005	0,135	0,208	-0,025		0,91
Faecalitalea contorta DSM3982, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	2,52	M110 K1	18,900	0,000	0,170	0,000	-0,440	0,000		0,92
Faecalitalea contorta DSM3982, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,23	M78	15,700	-0,245	-0,055	-0,415	0,218	-0,025		1,00
Faecalitalea contorta DSM3982, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,92	M110	NA	NA	NA	NA	NA	NA		1,10
Faecalitalea cylindroides DSM3983, Type strain	No	Firmicutes	Clostridiales	Faecalitalea	0,2	LYBHI 4 RU	NA	NA	NA	NA	NA	NA		1,01
Faecalitalea cylindroides DSM3983, Type strain	No	Firmicutes	Clostridiales	Faecalitalea	2,21	M110	1,990	0,000	0,000	0,000	-0,440	0,000		1,08
Faecalitalea cylindroides DSM3983, Type strain	No	Firmicutes	Clostridiales	Faecalitalea	2,51	M104	4,050	0,000	0,000	0,000	0,000	0,000		1,10
Faecalitalea cylindroides DSM3983, Type strain	No	Firmicutes	Clostridiales	Faecalitalea	1,80	LYBHI 4 RU	2,210	-0,080	0,000	-0,040	0,050	0,000		1,15
Faecalitalea cylindroides DSM3983, Type strain	No	Firmicutes	Clostridiales	Faecalitalea	2,21	M110	1,990	0,000	0,000	0,000	-0,440	0,000		1,21
Faecalitalea cylindroides DSM3983, Type strain	No	Firmicutes	Clostridiales	Faecalitalea	0,35	M78	NA	NA	NA	NA	NA	NA		2,13
Faecalitalea fissicatena DSM3598, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,30	LYBHI 4 RU	11,420	-0,150	0,000	-0,060	-0,040	0,040		0,90
Faecalitalea fissicatena DSM3598, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,30	Wilkins-Chalgren Anaerobe Broth	18,170	0,000	0,000	0,000	0,020	0,000		1,21
Faecalitalea fissicatena DSM3598, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	0,26	M104 C	3,580	0,000	0,000	0,000	0,000	0,000		1,85
Flavonifractor plautii DSM6740	No	Firmicutes	Clostridiales	Flavonifractor	0,50	M78	-2,290	0,665	-0,045	2,915	-0,043	-0,025		0,90
Flavonifractor plautii DSM6740	No	Firmicutes	Clostridiales	Flavonifractor	0,86	M110	-1,360	0,795	-0,045	4,305	0,238	-0,025		1,12
Flavonifractor plautii DSM6740	No	Firmicutes	Clostridiales	Flavonifractor	NA	M110	4,470	9,865	-0,075	5,435	0,628	-0,025		1,12
Flavonifractor plautii DSM6740	No	Firmicutes	Clostridiales	Flavonifractor	NA	LYBHI 4 RU	15,990	13,280	0,010	10,660	0,110	0,050		1,17
Flavonifractor plautii DSM6740	No	Firmicutes	Clostridiales	Flavonifractor	1,47	M78	20,350	28,380	0,050	5,970	1,280	0,130		1,37
Fusobacterium naviforme DSM20699	No	Fusobacteria	Fusobacteriales	Fusobacterium	NA	M104	7,450	1,810	-1,610	23,660	0,000	0,000		2,11
Fusobacterium nucleatum subsp. nucleatum DSM 15643	No	Fusobacteria	Fusobacteriales	Fusobacterium	1,47	LYBHI 4 RU	NA	NA	NA	NA	NA	NA		1,41
Fusobacterium nucleatum subsp. vincentii DSM 19508	No	Fusobacteria	Fusobacteriales	Fusobacterium	NA	M104	7,800	2,060	0,000	16,650	0,030	0,010		4,79
Fusobacterium nucleatum subsp. vincentii DSM 19508	No	Fusobacteria	Fusobacteriales	Fusobacterium	0,40	M78	28,700	0,900	0,070	24,340	0,000	0,000		1,66
Hungatella hathewayi DSM 13479, Type strain	No	Firmicutes	Clostridiales	Hungatella	1,43	M104	NA	NA	NA	NA	NA	NA		0,69
Hungatella hathewayi DSM 13479, Type strain	No	Firmicutes	Clostridiales	Hungatella	0,59	M104	3,370	-0,170	-0,030	-0,040	-0,030	0,000		0,82
Hungatella hathewayi DSM 13479, Type strain	No	Firmicutes	Clostridiales	Hungatella	1,64	LYBHI 4 RU	20,840	1,690	0,080	0,870	-0,040	-0,020		1,03
Lachnoclostridium asparagiforme DSM 15981, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	2,16	LYBHI 4 RU	18,010	-0,120	0,000	-0,100	-0,010	0,000		0,74
Lachnoclostridium asparagiforme DSM 15981, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,14	M104	NA	NA	NA	NA	NA	NA		0,75
Lachnoclostridium asparagiforme DSM 15981, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,94	LYBHI 4 RU	18,830	-0,080	0,000	0,000	-0,090	0,000		0,76
Lachnoclostridium asparagiforme DSM 15981, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	0,32	M104	NA	NA	NA	NA	NA	NA		0,77
Lachnoclostridium asparagiforme DSM 15981, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,88	LYBHI 4	17,560	0,000	0,000	0,040	0,070	0,000		0,88
Lachnoclostridium asparagiforme DSM 15981, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,63	M104 G	NA	NA	NA	NA	NA	NA		0,89
Lachnoclostridium asparagiforme DSM 15981, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	0,51	M78 G	-4,100	-0,335	-0,075	-0,415	-0,073	-0,025		0,95
Lachnoclostridium asparagiforme DSM 15981, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	NA	M104	NA	NA	NA	NA	NA	NA		1,00
Lachnoclostridium asparagiforme DSM 15981, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,63	M104	16,290	0,000	0,000	0,000	0,000	0,000		1,03
Lachnoclostridium asparagiforme DSM 15981, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	0,30	M104	2,060	-0,070	-0,020	-0,030	0,160	0,000		1,16
Lachnoclostridium asparagiforme DSM 15981, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	2,00	Wilkins-Chalgren Anaerobe Broth	22,480	0,000	0,000	0,000	0,000	0,000		1,33
Lachnoclostridium bolteae DSM 15670, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,24	M104	16,160	-0,110	-0,080	-0,180	-0,030	0,000		0,87
Lachnoclostridium bolteae DSM 15670, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,67	M104	17,080	-0,170	-0,080	-0,180	-0,030	0,000		1,10
Lachnoclostridium bolteae DSM 15670, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,61	LYBHI 4 RU	21,900	0,370	0,000	0,170	-0,010	0,030		1,23
Lachnoclostridium clostridioforme DSM933	No	Firmicutes	Clostridiales	Lachnoclostridium	0,34	LYBHI 4 RU	19,230	-0,060	0,000	0,030	-0,060	0,000		0,93
Lachnoclostridium clostridioforme DSM933	No	Firmicutes	Clostridiales	Lachnoclostridium	1,20	M412	11,200	0,000	0,000	0,000	0,000	0,000		0,98
Lachnoclostridium indolis DSM755, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,18	M78	0,020	-0,305	-0,045	-0,255	-0,053	-0,025		0,84
Lachnoclostridium indolis DSM755, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,98	M110	15,340	-0,265	-0,035	0,305	0,278	-0,025		0,84
Lachnoclostridium indolis DSM755, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	0,90	M78	3,660	-0,335	-0,035	-0,235	0,007	-0,025		0,90
Lachnoclostridium indolis DSM755, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,72	LYBHI 4 RU	16,070	0,510	0,030	0,250	0,070	0,000		1,22
Lachnoclostridium lavalense DSM 19851, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,70	LYBHI 4 RU	12,830	-0,050	0,000	0,010	-0,030	0,040		0,81
Lachnoclostridium lavalense DSM 19851, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,49	M215c	11,750	0,000	0,000	0,000	-0,030	0,000		0,83
Lachnoclostridium phytofermentans DSM 18823, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	0,46	LYBHI 4 RU	NA	NA	NA	NA	NA	NA		0,22
Lachnoclostridium phytofermentans DSM 18823, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,08	M110	NA	NA	NA	NA	NA	NA		0,36

Lachnoclostridium phytofermentans DSM18823, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	0,5	Wilkins-Chalgren Anaerobe Broth	NA	NA	NA	NA	NA	NA	0,40
Lachnoclostridium phytofermentans DSM18823, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,08	M78	NA	NA	NA	NA	NA	NA	1,20
Lachnoclostridium saccharolyticum DSM2544, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,99	M110	8,230	-0,295	-0,075	-0,275	0,218	-0,025	0,77
Lachnoclostridium saccharolyticum DSM2544, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,20	M110	9,270	-0,315	-0,075	-0,295	0,498	-0,025	1,02
Lachnoclostridium saccharolyticum DSM2544, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,49	M78	1,510	-0,325	-0,035	-0,415	-0,173	-0,025	1,04
Lachnoclostridium scindens DSM5676, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,38	LYBHI 4 RU	20,240	-0,100	0,000	0,010	-0,010	0,040	0,95
Lachnoclostridium scindens DSM5676, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	NA	M110	NA	NA	NA	NA	NA	NA	0,96
Lachnoclostridium sphenoides DSM632, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,98	M110	10,910	-0,51	-0,08	-0,07	0,15	-0,03	1,07
Lachnoclostridium sphenoides DSM632, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	NA	M78	NA	NA	NA	NA	NA	NA	1,17
Lachnoclostridium symbiosum DSM934, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,69	M110	11,630	-0,15	-0,08	6,11	0,24	-0,03	1,44
Lachnoclostridium symbiosum DSM934, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,14	LYBHI 4 RU	24,650	0,11	0,02	10,73	0,06	-0,12	1,51
Lachnoclostridium symbiosum DSM934, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	0,64	M78	3,220	-0,25	-0,04	6,99	-0,04	-0,03	1,58
Lachnoclostridium symbiosum DSM934, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	0,52	M78	6,050	-0,30	-0,08	8,15	-0,05	-0,03	1,71
Lachnoclostridium symbiosum DSM934, Type strain	No	Firmicutes	Clostridiales	Lachnoclostridium	1,23	M110	22,710	0,16	-0,04	14,70	0,29	-0,03	1,78
Lactobacillus acidophilus DSM20079, Type strain	No	Firmicutes	Lactobacillales	Lactobacillus	2,00	M58	NA	NA	NA	NA	NA	NA	1,95
Lactobacillus acidophilus DSM20079, Type strain	No	Firmicutes	Lactobacillales	Lactobacillus	2,00	M58	NA	NA	NA	NA	NA	NA	2,02
Lactobacillus acidophilus DSM20079, Type strain	No	Firmicutes	Lactobacillales	Lactobacillus	1,71	M58	NA	NA	NA	NA	NA	NA	2,48
Lactobacillus casei, Z313	No	Firmicutes	Lactobacillales	Lactobacillus	2,00	M58	NA	NA	NA	NA	NA	NA	1,37
Lactobacillus casei, Z313	No	Firmicutes	Lactobacillales	Lactobacillus	2,00	M58	NA	NA	NA	NA	NA	NA	1,40
Lactobacillus johnsonii, La-1	No	Firmicutes	Lactobacillales	Lactobacillus	NA	M58	NA	NA	NA	NA	NA	NA	1,69
Lactobacillus johnsonii, Ncc 533	No	Firmicutes	Lactobacillales	Lactobacillus	1,72	M58	NA	NA	NA	NA	NA	NA	2,60
Lactobacillus mucosae DSM13345	No	Firmicutes	Lactobacillales	Lactobacillus	0,86	LYBHI 4	NA	NA	NA	NA	NA	NA	1,33
Lactobacillus paracasei, Z763	No	Firmicutes	Lactobacillales	Lactobacillus	2,00	M58	NA	NA	NA	NA	NA	NA	1,62
Lactobacillus paracasei, Z763	No	Firmicutes	Lactobacillales	Lactobacillus	2,00	M58	NA	NA	NA	NA	NA	NA	1,90
Lactobacillus plantarum ATCC43199, Type strain	No	Firmicutes	Lactobacillales	Lactobacillus	0,67	M58	NA	NA	NA	NA	NA	NA	1,09
Lactobacillus plantarum ATCC43199, Type strain	No	Firmicutes	Lactobacillales	Lactobacillus	0,71	M58	NA	NA	NA	NA	NA	NA	1,09
Lactobacillus reuteri DSM17509	No	Firmicutes	Lactobacillales	Lactobacillus	2,00	M58	NA	NA	NA	NA	NA	NA	1,99
Lactobacillus reuteri DSM17509	No	Firmicutes	Lactobacillales	Lactobacillus	2,00	M58	NA	NA	NA	NA	NA	NA	2,16
Lactobacillus reuteri DSM20016, Type strain	No	Firmicutes	Lactobacillales	Lactobacillus	1,22	M58	NA	NA	NA	NA	NA	NA	2,61
Lactobacillus reuteri DSM20056, Type strain	No	Firmicutes	Lactobacillales	Lactobacillus	2,00	M58	NA	NA	NA	NA	NA	NA	1,92
Lactobacillus reuteri DSM20056, Type strain	No	Firmicutes	Lactobacillales	Lactobacillus	2,00	M58	NA	NA	NA	NA	NA	NA	2,39
Lactobacillus rhamnosus, Z712	No	Firmicutes	Lactobacillales	Lactobacillus	1,57	M58	NA	NA	NA	NA	NA	NA	3,35
Lactobacillus rhamnosus, Z712	No	Firmicutes	Lactobacillales	Lactobacillus	2,00	M58	NA	NA	NA	NA	NA	NA	1,53
Lactobacillus rhamnosus, Z712	No	Firmicutes	Lactobacillales	Lactobacillus	2,00	M58	NA	NA	NA	NA	NA	NA	1,71
Marvinbryantia formatexigens DSM14469, Type strain	No	Firmicutes	Clostridiales	Marvinbryantia	0,59	LYBHI 4	NA	NA	NA	NA	NA	NA	1,22
Marvinbryantia formatexigens DSM14469, Type strain	No	Firmicutes	Clostridiales	Marvinbryantia	0,28	M110	NA	NA	NA	NA	NA	NA	1,30
Marvinbryantia formatexigens DSM14469, Type strain	No	Firmicutes	Clostridiales	Marvinbryantia	0,31	LYBHI 4	NA	NA	NA	NA	NA	NA	1,48
Marvinbryantia formatexigens DSM14469, Type strain	No	Firmicutes	Clostridiales	Marvinbryantia	0,31	LYBHI 4 RU	98,560	26,27	3,03	14,22	5,68	2,68	1,51
Marvinbryantia formatexigens DSM14469, Type strain	No	Firmicutes	Clostridiales	Marvinbryantia	0,67	LYBHI 4 RU	NA	NA	NA	NA	NA	NA	1,57
Oscillibacter valericigenes DSM18026, Type strain	No	Firmicutes	Clostridiales	Oscillibacter	1,04	LYBHI 4 RU	-5,080	-0,120	0,510	10,890	1,210	0,000	1,14
Oscillibacter valericigenes DSM18026, Type strain	No	Firmicutes	Clostridiales	Oscillibacter	1,05	LYBHI 4 RU	NA	NA	NA	NA	NA	NA	2,15
Oscillibacter valericigenes DSM18026, Type strain	No	Firmicutes	Clostridiales	Oscillibacter	0,75	M104	6,010	0,94	0,17	11,28	0,70	0,00	2,32
Paeniclostridium ordellii DSM2141, Type strain	No	Firmicutes	Clostridiales	Peptoclostridium	0,14	LYBHI 4	4,380	0,000	-0,130	0,050	0,000	0,000	1,17
Parabacteroides distasonis ATCC8503, Type strain	No	Bacteroidetes	Bacteroidales	Parabacteroides	0,67	LYHBHI 1	NA	NA	NA	NA	NA	NA	1,13
Parabacteroides distasonis ATCC8503, Type strain	No	Bacteroidetes	Bacteroidales	Parabacteroides	0,67	LYHBHI 1	9,520	32,770	0,830	-0,220	5,570	0,000	1,66
Parabacteroides johnsonii DSM18315, Type strain	No	Bacteroidetes	Bacteroidales	Parabacteroides	0,60	M104	6,995	2,260	-1,790	-0,240	0,350	0,000	1,23
Prevotella albensis DSM11370, Type strain	No	Bacteroidetes	Bacteroidales	Prevotella	1,12	M78	NA	NA	NA	NA	NA	NA	0,97
Prevotella albensis DSM11370, Type strain	No	Bacteroidetes	Bacteroidales	Prevotella	1,32	Wilkins-Chalgren Anaerobe Broth	NA	NA	NA	NA	NA	NA	1,25
Prevotella bivia DSM20514, Type strain	No	Bacteroidetes	Bacteroidales	Prevotella	1,73	M104	NA	NA	NA	NA	NA	NA	0,99
Prevotella copri DSM18205, Type strain	No	Bacteroidetes	Bacteroidales	Prevotella	0,73	M104	NA	NA	NA	NA	NA	NA	1,10
Prevotella copri DSM18205, Type strain	No	Bacteroidetes	Bacteroidales	Prevotella	0,44	LYBHI 4	NA	NA	NA	NA	NA	NA	1,16
Prevotella copri DSM18205, Type strain	No	Bacteroidetes	Bacteroidales	Prevotella	0,44	LYBHI 4	NA	NA	NA	NA	NA	NA	1,21
Prevotella copri DSM18205, Type strain	No	Bacteroidetes	Bacteroidales	Prevotella	1,17	M104	NA	NA	NA	NA	NA	NA	1,25
Prevotella copri DSM18205, Type strain	No	Bacteroidetes	Bacteroidales	Prevotella	0,89	LYBHI 4	NA	NA	NA	NA	NA	NA	1,26
Prevotella copri DSM18205, Type strain	No	Bacteroidetes	Bacteroidales	Prevotella	0,51	LYBHI 4 RU	NA	NA	NA	NA	NA	NA	1,28
Prevotella copri DSM18205, Type strain	No	Bacteroidetes	Bacteroidales	Prevotella	0,44	M104	21,740	0,000	-1,210	-0,190	0,000	0,000	2,89
Prevotella oris DSM18711, Type strain	No	Bacteroidetes	Bacteroidales	Prevotella	1,66	M104	NA	NA	NA	NA	NA	NA	1,45
Propionibacterium acnes, B15	No	Actinobacteria	Propionibacteriales	Propionibacterium	0,84	M104	-0,170	0,690	-1,490	-0,160	0,020	0,000	1,30
Roseburia faecis DSM16840, Type strain	No	Firmicutes	Clostridiales	Roseburia	0,62	LYHBHI 1	1,610	0,000	-0,200	0,020	0,000	0,000	1,64
Roseburia hominis DSM16839, Type strain	No	Firmicutes	Clostridiales	Roseburia	0,68	M58	-0,690	0,000	-0,280	10,370	0,000	0,000	2,39
Roseburia hominis DSM16839, Type strain	No	Firmicutes	Clostridiales	Roseburia	0,72	M58	-11,720	-0,140	-0,030	11,620	-0,010	0,000	2,96
Roseburia intestinalis DSM14610, Type strain	No	Firmicutes	Clostridiales	Roseburia	1,09	LYBHI 4 RU	NA	NA	NA	NA	NA	NA	1,49
Roseburia intestinalis DSM14610, Type strain	No	Firmicutes	Clostridiales	Roseburia	0,68	LYBHI 4	-2,510	0,010	-0,030	13,230	-0,020	0,000	1,52
Roseburia intestinalis DSM14610, Type strain	No	Firmicutes	Clostridiales	Roseburia	0,86	M104	-1,100	2,850	0,410	30,230	0,890	0,680	1,54
Roseburia intestinalis DSM14610, Type strain	No	Firmicutes	Clostridiales	Roseburia	0,39	LYBHI 4	-2,290	-0,010	-0,030	11,690	-0,470	0,000	1,75
Roseburia intestinalis DSM14610, Type strain	No	Firmicutes	Clostridiales	Roseburia	0,80	M104	-1,840	0,030	-0,080	8,030	-0,030	0,000	2,34

Roseburia intestinalis DSM14610, Type strain	No	Firmicutes	Clostridiales	Roseburia	0,22	M58	-1,880	0,000	-0,030	12,990	0,000	0,000	2,64
Roseburia intestinalis DSM14610, Type strain	No	Firmicutes	Clostridiales	Roseburia	0,39	LYBHI 4	-2,290	-0,010	-0,030	11,690	-0,470	0,000	2,66
Roseburia intestinalis DSM14610, Type strain	No	Firmicutes	Clostridiales	Roseburia	0,68	LYBHI 4	-2,510	0,010	-0,030	13,230	-0,020	0,000	2,88
Roseburia intestinalis DSM14610, Type strain	No	Firmicutes	Clostridiales	Roseburia	NA	M58	-15,270	-0,170	-0,010	21,480	-0,040	0,000	3,84
Roseburia inulinivorans DSM16841, Type strain	No	Firmicutes	Clostridiales	Roseburia	0,46	LYBHI 4	106,610	28,980	3,200	15,590	5,990	2,900	1,22
Roseburia inulinivorans DSM16841, Type strain	No	Firmicutes	Clostridiales	Roseburia	0,90	LYBHI 4 RJ	13,420	5,480	0,350	1,890	0,690	0,300	1,29
Roseburia inulinivorans DSM16841, Type strain	No	Firmicutes	Clostridiales	Roseburia	0,94	M412	-0,350	0,000	0,000	3,490	0,000	0,000	1,38
Roseburia inulinivorans DSM16841, Type strain	No	Firmicutes	Clostridiales	Roseburia	NA	M104	-4,610	0,000	0,000	4,180	0,000	0,000	2,90
Ruminoclostridium leptum DSM753, Type strain	No	Firmicutes	Clostridiales	Ruminiclostridium	0,56	LYBHI 4 RJ	-2,120	-1,320	-0,090	-0,870	-0,190	-0,120	0,75
Ruminoclostridium leptum DSM753, Type strain	No	Firmicutes	Clostridiales	Ruminiclostridium	NA	LYBHI 4 RJ	NA	NA	NA	NA	NA	NA	0,77
Ruminoclostridium leptum DSM753, Type strain	No	Firmicutes	Clostridiales	Ruminiclostridium	NA	LYBHI 4	3,920	0,200	-0,010	0,060	0,070	0,000	0,80
Ruminoclostridium leptum DSM753, Type strain	No	Firmicutes	Clostridiales	Ruminiclostridium	0,63	LYBHI 4 RJ	4,170	0,180	-0,020	-0,160	-0,080	0,000	0,81
Ruminoclostridium leptum DSM753, Type strain	No	Firmicutes	Clostridiales	Ruminiclostridium	0,56	LYBHI 4	4,530	0,100	-0,040	0,050	-0,470	0,000	0,89
Ruminoclostridium leptum DSM753, Type strain	No	Firmicutes	Clostridiales	Ruminiclostridium	0,22	LYHBHI 1	6,420	0,100	0,030	0,070	0,020	0,000	0,91
Ruminoclostridium leptum DSM753, Type strain	No	Firmicutes	Clostridiales	Ruminiclostridium	0,22	LYHBHI 1	4,230	0,100	-0,170	-0,170	0,020	0,000	1,11
Ruminoclostridium leptum DSM753, Type strain	No	Firmicutes	Clostridiales	Ruminiclostridium	0,56	LYBHI 4 RJ	6,300	0,480	0,010	-0,020	0,000	-0,120	1,20
Ruminoclostridium siraeum DSM3996	No	Firmicutes	Clostridiales	Ruminiclostridium	NA	M104 RJ	NA	NA	NA	NA	NA	NA	0,98
Ruminoclostridium siraeum DSM3996	No	Firmicutes	Clostridiales	Ruminiclostridium	0,66	M110 RJ	3,480	0,000	0,060	0,010	0,030	0,000	1,03
Ruminoclostridium siraeum DSM3996	No	Firmicutes	Clostridiales	Ruminiclostridium	0,70	LYBHI 4 RJ	NA	NA	NA	NA	NA	NA	1,17
Ruminoclostridium siraeum DSM3996	No	Firmicutes	Clostridiales	Ruminiclostridium	0,46	LYBHI 4 RJ	9,260	1,550	0,000	1,010	0,120	0,060	1,18
Ruminoclostridium siraeum DSM3996	No	Firmicutes	Clostridiales	Ruminiclostridium	0,85	M110 RJ	-3,480	-1,430	0,000	-0,840	-0,300	-0,120	1,25
Ruminoclostridium siraeum DSM3996	No	Firmicutes	Clostridiales	Ruminiclostridium	NA	M104	NA	NA	NA	NA	NA	NA	1,52
Ruminoclostridium sporosphaeroides DSM1294, Type strain	No	Firmicutes	Clostridiales	Ruminiclostridium	1,85	LYBHI 1	19,770	0,130	0,050	0,200	0,100	NA	1,00
Ruminoclostridium sporosphaeroides DSM1294, Type strain	No	Firmicutes	Clostridiales	Ruminiclostridium	1,09	LYBHI 1	10,670	0,330	0,000	2,310	-0,060	0,000	1,01
Ruminoclostridium sporosphaeroides DSM1294, Type strain	No	Firmicutes	Clostridiales	Ruminiclostridium	0,97	M78	16,170	11,665	-0,045	0,765	-0,093	-0,025	1,12
Ruminoclostridium sporosphaeroides DSM1294, Type strain	No	Firmicutes	Clostridiales	Ruminiclostridium	1,85	M78	3,940	-0,405	-0,075	0,165	-0,053	-0,025	1,13
Ruminoclostridium sporosphaeroides DSM1294, Type strain	No	Firmicutes	Clostridiales	Ruminiclostridium	0,12	LYBHI 1	26,070	0,900	0,000	6,450	-0,070	0,000	1,51
Ruminoclostridium sporosphaeroides DSM1294, Type strain	No	Firmicutes	Clostridiales	Ruminiclostridium	1,02	LYBHI 1	27,530	1,280	0,000	7,480	-0,070	0,000	1,58
Ruminoclostridium sporosphaeroides DSM1294, Type strain	No	Firmicutes	Clostridiales	Ruminiclostridium	1,06	M78	NA	NA	NA	NA	NA	NA	1,61
Ruminoclostridium sporosphaeroides DSM1294, Type strain	No	Firmicutes	Clostridiales	Ruminiclostridium	1,35	LYBHI 4 RJ	NA	NA	NA	NA	NA	NA	1,72
Ruminoclostridium sporosphaeroides DSM1294, Type strain	No	Firmicutes	Clostridiales	Ruminiclostridium	0,46	M78	50,230	18,960	0,060	4,750	0,180	0,000	1,80
Ruminoclostridium sporosphaeroides DSM1294, Type strain	No	Firmicutes	Clostridiales	Ruminiclostridium	1,48	M110	NA	NA	NA	NA	NA	NA	1,90
Ruminococcus bromii ATCC27255, Type strain	No	Firmicutes	Clostridiales	Ruminococcus	1,96	LYBHI 4	1,950	0,140	0,020	0,190	0,100	0,000	0,86
Ruminococcus bromii ATCC27255, Type strain	No	Firmicutes	Clostridiales	Ruminococcus	1,92	LYBHI 4 RJ	-0,640	0,340	0,030	0,280	0,010	0,000	1,10
Ruminococcus bromii ATCC27255, Type strain	No	Firmicutes	Clostridiales	Ruminococcus	1,98	LYBHI 4	4,390	0,290	0,030	0,200	0,080	0,000	1,18
Ruminococcus bromii ATCC27255, Type strain	No	Firmicutes	Clostridiales	Ruminococcus	NA	LYBHI 4	NA	NA	NA	NA	NA	NA	1,37
Ruminococcus lactaris ATCC29176, Type strain	No	Firmicutes	Clostridiales	Ruminococcus	1,48	LYBHI 4	9,410	0,420	0,050	0,290	-0,280	0,000	0,93
Ruminococcus lactaris ATCC29176, Type strain	No	Firmicutes	Clostridiales	Ruminococcus	NA	LYBHI 4	3,710	0,100	0,000	0,060	0,010	0,00	0,97
Ruminococcus lactaris ATCC29176, Type strain	No	Firmicutes	Clostridiales	Ruminococcus	1,40	LYBHI 1	9,900	0,280	0,050	0,220	0,080	0,000	1,00
Ruminococcus lactaris ATCC29176, Type strain	No	Firmicutes	Clostridiales	Ruminococcus	0,97	LYBHI 4 RJ	7,740	0,510	0,060	0,280	0,160	0,01	1,03
Ruminococcus lactaris ATCC29176, Type strain	No	Firmicutes	Clostridiales	Ruminococcus	0,49	LYBHI 4	4,830	0,400	0,060	0,290	0,170	0,00	1,03
Ruminococcus lactaris ATCC29176, Type strain	No	Firmicutes	Clostridiales	Ruminococcus	0,94	LYBHI 4	4,480	0,210	0,020	0,120	0,050	0,00	1,05
Ruminococcus torques ATCC27756, Type strain	No	Firmicutes	Clostridiales	Ruminococcus	0,32	M104	5,150	0,000	-1,360	-0,190	0,000	0,00	1,42
Sutterella parivubra DSM19354, Type strain	No	Proteobacteria	Betaproteobacteria	Sutterellaceae	0,62	LYBHI 4	NA	NA	NA	NA	NA	NA	3,14
Sutterella parivubra DSM19354, Type strain	No	Proteobacteria	Betaproteobacteria	Sutterellaceae	0,96	LYBHI 1	NA	NA	NA	NA	NA	NA	3,27
Sutterella wadsworthensis DSM14016, Type strain	No	Proteobacteria	Betaproteobacteria	Sutterellaceae	0,37	LYBHI 4 mucin	NA	NA	NA	NA	NA	NA	2,41
Sutterella wadsworthensis DSM14016, Type strain	No	Proteobacteria	Betaproteobacteria	Sutterellaceae	NA	M110 RJ	NA	NA	NA	NA	NA	NA	2,82
Sutterella wadsworthensis DSM14016, Type strain	No	Proteobacteria	Betaproteobacteria	Sutterellaceae	0,07	LYBHI 4 RJ	NA	NA	NA	NA	NA	NA	2,84
Tyzzerella colinum DSM6011, Type strain	No	Firmicutes	Clostridiales	Tyzzerella	0,70	M78	2,880	-0,415	-0,075	-0,415	-0,073	-0,025	1,17
Tyzzerella colinum DSM6011, Type strain	No	Firmicutes	Clostridiales	Tyzzerella	0,39	M78	-1,220	-0,355	-0,075	-0,415	-0,033	-0,025	1,18
Tyzzerella nexilis DSM1787, Type strain	No	Firmicutes	Clostridiales	Tyzzerella	0,65	M110	2,800	-0,355	-0,045	-0,275	0,738	-0,03	0,85
Tyzzerella nexilis DSM1787, Type strain	No	Firmicutes	Clostridiales	Tyzzerella	1,34	M110	4,010	-0,285	-0,045	-0,415	0,308	-0,03	0,86
Tyzzerella nexilis DSM1787, Type strain	No	Firmicutes	Clostridiales	Tyzzerella	0,30	LYBHI	6,170	0,000	-0,150	-0,200	0,000	0,00	1,05
Tyzzerella nexilis DSM1787, Type strain	No	Firmicutes	Clostridiales	Tyzzerella	0,37	M78	-5,330	-0,315	-0,075	-0,255	0,007	-0,03	1,07

Marinelli *et al.* Supplementary Table 2

Short name	Name	Origin	Composition
LB	Luria-Bertani broth		
BHI aerobic	Brain Heart Infusion	BD Difco	
LYBHI	Liquid Yeast Brain Infusion	Homemade	BHI 37g/L+ Yeast extact 5g/ml+ Hemin 0.1% 10mL/L+ Resazurin 0.1% 1ml/L
LYBHI 1	Liquid Yeast Brain Heart Infusion 1	Homemade	LYBHI+ Cystein HCL 0.6g/L
LYBHI 1 mucin	Liquid Yeast Brain Heart Infusion 1 mucin	Homemade	LYBHI+ Cystein HCL 0.6g/L+ mucin type II 1g/L
LYBHI 2	Liquid Yeast Brain Heart Infusion 2	Homemade	LYBHI+ Cystein HCL 0.6g/L+ cellobiose 0.5g/L+ maltose 0.5g/L
LYBHI 3	Liquid Yeast Brain Heart Infusion 3	Homemade	LYBHI+ Cystein HCL 0.6g/L+ cellobiose 0.5g/L+ maltose 0.5g/L + K1 Vitamin
LYBHI 4	Liquid Yeast Brain Heart Infusion 4	Homemade	LYBHI+ Cystein HCL 0.6g/L+ cellobiose 0.5g/L+ maltose 0.5g/L+ soluble starch 0.5g/L
LYBHI 4 Mucin	Liquid Yeast Brain Heart Infusion 4 Mucin	Homemade	LYBHI+ Cystein HCL 0.6g/L+ cellobiose 0.5g/L+ maltose 0.5g/L+ soluble starch 0.5g/L+ mucin type II 1g/L
LYBHI 4 CH/PL	Liquid Yeast Brain Heart Infusion 4 Cholesterol Phospholipids	Homemade	LYBHI+ Cystein HCL 0.6g/L+ cellobiose 0.5g/L+ maltose 0.5g/L+ starch 0.5g/L+ cholesterol 2g/L + phopholides 1g/L
LYBHI 4 RJ	Liquid Yeast Brain Heart Infusion 4 Rumen Juice	Homemade	LYBHI+ Cystein HCL 0.6g/L+ cellobiose 0.5g/L+ maltose 0.5g/L+ starch 0.5g/L+ rumen fluid 10% (v/v)
M1	DSMZ NUTRIENT AGAR	Homemade	
M104	DSMZ PYG MEDIUM + K1 Vitamin	Homemade	
M104c	DSMZ PY + X MEDIUM	Homemade	
M110	DSMZ CHOPPED MEAT MEDIUM WITH CARBOHYDRATES	Homemade	
M110 K1	DSMZ CHOPPED MEAT MEDIUM WITH CARBOHYDRATES WITH VITK1	Homemade	
M110 RJ	DSMZ CHOPPED MEAT MEDIUM WITH CARBOHYDRATES WITH RUMEN JUCE	Homemade	
M215c	BHI MEDIUM FOR STRICT ANAEROBES	Homemade	
M3	DSMZ AZOTOBACTER MEDIUM	Homemade	
M412	DSMZ ACETOMICROBIUM FAECALIS MEDIUM	Homemade	
M436	DSMZ RUMINOCOCCUS ALBUS MEDIUM	Homemade	
M58	DSMZ BIFIDOBACTERIUM MEDIUM	Homemade	
M78	DSMZ CHOPPED MEAT MEDIUM	Homemade	
M339	DSMZ WILKINS-CHALGREN ANAEROBE BROTH	Homemade	

References

1 Postler, T. S. & Ghosh, S. Understanding the Holobiont: How Microbial Metabolites Affect Human Health and Shape the Immune System. *Cell metabolism* **26**, 110-130, doi:10.1016/j.cmet.2017.05.008 (2017).

2 Gao, J. *et al.* Impact of the Gut Microbiota on Intestinal Immunity Mediated by Tryptophan Metabolism. *Front Cell Infect Microbiol* **8**, 13, doi:10.3389/fcimb.2018.00013 (2018).

3 Quintana, F. J. & Sherr, D. H. Aryl hydrocarbon receptor control of adaptive immunity. *Pharmacol Rev* **65**, 1148-1161, doi:10.1124/pr.113.007823 (2013).

4 Veldhoen, M. & Ferreira, C. Influence of nutrient-derived metabolites on lymphocyte immunity. *Nature medicine* **21**, 709-718, doi:10.1038/nm.3894 (2015).

5 Barouki, R., Aggerbeck, M., Aggerbeck, L. & Coumoul, X. The aryl hydrocarbon receptor system. *Drug Metabol Drug Interact* **27**, 3-8, doi:10.1515/dmdi-2011-0035 (2012).

6 Hubbard, T. D. *et al.* Adaptation of the human aryl hydrocarbon receptor to sense microbiota-derived indoles. *Scientific reports* **5**, 12689, doi:10.1038/srep12689 (2015).

7 Hubbard, T. D., Murray, I. A. & Perdew, G. H. Indole and Tryptophan Metabolism: Endogenous and Dietary Routes to Ah Receptor Activation. *Drug Metab Dispos* **43**, 1522-1535, doi:10.1124/dmd.115.064246 (2015).

8 Jin, U. H. *et al.* Microbiome-derived tryptophan metabolites and their aryl hydrocarbon receptor-dependent agonist and antagonist activities. *Mol Pharmacol* **85**, 777-788, doi:10.1124/mol.113.091165 (2014).

9 Fujii-Kuriyama, Y., Ema, M., Mimura, J. & Sogawa, K. Ah receptor: a novel ligand-activated transcription factor. *Exp Clin Immunogenet* **11**, 65-74 (1994).

10 Hankinson, O. The aryl hydrocarbon receptor complex. *Annu Rev Pharmacol Toxicol* **35**, 307-340, doi:10.1146/annurev.pa.35.040195.001515 (1995).

11 Lamas, B. *et al.* CARD9 impacts colitis by altering gut microbiota metabolism of tryptophan into aryl hydrocarbon receptor ligands. *Nature medicine* **22**, 598-605, doi:10.1038/nm.4102 (2016).

12 Monteleone, I. *et al.* Aryl hydrocarbon receptor-induced signals up-regulate IL-22 production and inhibit inflammation in the gastrointestinal tract. *Gastroenterology* **141**, 237-248, 248 e231, doi:10.1053/j.gastro.2011.04.007 (2011).

13 Furumatsu, K. *et al.* A role of the aryl hydrocarbon receptor in attenuation of colitis. *Dig Dis Sci* **56**, 2532-2544, doi:10.1007/s10620-011-1643-9 (2011).

14 Goettel, J. A. *et al.* AHR Activation Is Protective against Colitis Driven by T Cells in Humanized Mice. *Cell reports* **17**, 1318-1329, doi:10.1016/j.celrep.2016.09.082 (2016).

15 Liu, Z. *et al.* Aryl hydrocarbon receptor activation maintained the intestinal epithelial barrier function through Notch1 dependent signaling pathway. *Int J Mol Med* **41**, 1560-1572, doi:10.3892/ijmm.2017.3341 (2018).

16 Yu, M. *et al.* Aryl Hydrocarbon Receptor Activation Modulates Intestinal Epithelial Barrier Function by Maintaining Tight Junction Integrity. *Int J Biol Sci* **14**, 69-77, doi:10.7150/ijbs.22259 (2018).

17 Schiering, C., Vonk, A., Das, S., Stockinger, B. & Wincent, E. Cytochrome P4501-inhibiting chemicals amplify aryl hydrocarbon receptor activation and IL-22 production in T helper 17 cells. *Biochem Pharmacol* **151**, 47-58, doi:10.1016/j.bcp.2018.02.031 (2018).

18 Schiering, C. *et al.* Feedback control of AHR signalling regulates intestinal immunity. *Nature* **542**, 242-245, doi:10.1038/nature21080 (2017).

19 Zelante, T. *et al.* Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22. *Immunity* **39**, 372-385, doi:10.1016/j.immuni.2013.08.003 (2013).

20 Sonowal, R. *et al.* Indoles from commensal bacteria extend healthspan. *Proceedings of the National Academy of Sciences of the United States of America* **114**, E7506-E7515, doi:10.1073/pnas.1706464114 (2017).

21 Vital, M., Howe, A. C. & Tiedje, J. M. Revealing the bacterial butyrate synthesis pathways by analyzing (meta)genomic data. *mBio* **5**, e00889, doi:10.1128/mBio.00889-14 (2014).

22 Cummings, J. H. Short chain fatty acids in the human colon. *Gut* **22**, 763-779 (1981).

23 Cummings, J. H., Pomare, E. W., Branch, W. J., Naylor, C. P. & Macfarlane, G. T. Short chain fatty acids in human large intestine, portal, hepatic and venous blood. *Gut* **28**, 1221-1227 (1987).

24 Blachier, F., Mariotti, F., Huneau, J. F. & Tome, D. Effects of amino acid-derived luminal metabolites on the colonic epithelium and physiopathological consequences. *Amino Acids* **33**, 547-562, doi:10.1007/s00726-006-0477-9 (2007).

25 Schaible, U. E. & Kaufmann, S. H. A nutritive view on the host-pathogen interplay. *Trends in microbiology* **13**, 373-380, doi:10.1016/j.tim.2005.06.009 (2005).

26 Schilderink, R., Verseijden, C. & de Jonge, W. J. Dietary inhibitors of histone deacetylases in intestinal immunity and homeostasis. *Frontiers in immunology* **4**, 226, doi:10.3389/fimmu.2013.00226 (2013).

27 Candido, E. P., Reeves, R. & Davie, J. R. Sodium butyrate inhibits histone deacetylation in cultured cells. *Cell* **14**, 105-113 (1978).

28 Sealy, L. & Chalkley, R. The effect of sodium butyrate on histone modification. *Cell* **14**, 115-121 (1978).

29 Hinnebusch, B. F., Meng, S., Wu, J. T., Archer, S. Y. & Hodin, R. A. The effects of short-chain fatty acids on human colon cancer cell phenotype are associated with histone hyperacetylation. *J Nutr* **132**, 1012-1017 (2002).

30 Tolhurst, G. *et al.* Short-chain fatty acids stimulate glucagon-like peptide-1 secretion via the G-protein-coupled receptor FFAR2. *Diabetes* **61**, 364-371, doi:10.2337/db11-1019 (2012).

31 Cuff, M., Dyer, J., Jones, M. & Shirazi-Beechey, S. The human colonic monocarboxylate transporter Isoform 1: its potential importance to colonic tissue homeostasis. *Gastroenterology* **128**, 676-686 (2005).

32 Jin, U. H. *et al.* Short Chain Fatty Acids Enhance Aryl Hydrocarbon (Ah) Responsiveness in Mouse Colonocytes and Caco-2 Human Colon Cancer Cells. *Scientific reports* **7**, 10163, doi:10.1038/s41598-017-10824-x (2017).

33 Gallinari, P., Di Marco, S., Jones, P., Pallaoro, M. & Steinkuhler, C. HDACs, histone deacetylation and gene transcription: from molecular biology to cancer therapeutics. *Cell Res* **17**, 195-211, doi:10.1038/sj.cr.7310149 (2007).

34 Kinoshita, M., Suzuki, Y. & Saito, Y. Butyrate reduces colonic paracellular permeability by enhancing PPARgamma activation. *Biochem Biophys Res Commun* **293**, 827-831, doi:10.1016/S0006-291X(02)00294-2 (2002).

35 Alex, S. *et al.* Short-chain fatty acids stimulate angiopoietin-like 4 synthesis in human colon adenocarcinoma cells by activating peroxisome proliferator-activated receptor gamma. *Mol Cell Biol* **33**, 1303-1316, doi:10.1128/MCB.00858-12 (2013).

36 Yin, Z., Henry, E. C. & Gasiewicz, T. A. (-)-Epigallocatechin-3-gallate is a novel Hsp90 inhibitor. *Biochemistry* **48**, 336-345, doi:10.1021/bi801637q (2009).

37 Perkins, A. *et al.* A Structural Switch between Agonist and Antagonist Bound Conformations for a Ligand-Optimized Model of the Human Aryl Hydrocarbon Receptor Ligand Binding Domain. *Biology (Basel)* **3**, 645-669, doi:10.3390/biology3040645 (2014).

38 Smith, K. J. *et al.* Identification of a high-affinity ligand that exhibits complete aryl hydrocarbon receptor antagonism. *J Pharmacol Exp Ther* **338**, 318-327, doi:10.1124/jpet.110.178392 (2011).

39 Palermo, C. M., Westlake, C. A. & Gasiewicz, T. A. Epigallocatechin gallate inhibits aryl hydrocarbon receptor gene transcription through an indirect mechanism involving binding to a 90 kDa heat shock protein. *Biochemistry* **44**, 5041-5052, doi:10.1021/bi047433p (2005).

40 Henry, E. C. & Gasiewicz, T. A. Agonist but not antagonist ligands induce conformational change in the mouse aryl hydrocarbon receptor as detected by partial proteolysis. *Mol Pharmacol* **63**, 392-400 (2003).

41 Bonati, L., Corrada, D., Tagliabue, S. G. & Motta, S. Molecular modeling of the AhR structure and interactions can shed light on ligand-dependent activation and transformation mechanisms. *Curr Opin Toxicol* **2**, 42-49, doi:10.1016/j.cotox.2017.01.011 (2017).

42 Bisson, W. H. *et al.* Modeling of the aryl hydrocarbon receptor (AhR) ligand binding domain and its utility in virtual ligand screening to predict new AhR ligands. *J Med Chem* **52**, 5635-5641, doi:10.1021/jm900199u (2009).

43 Erbel, P. J. *et al.* Solution structure of the alpha-subunit of human chorionic gonadotropin. *Eur J Biochem* **260**, 490-498 (1999).

44 Stockinger, B., Di Meglio, P., Gialitakis, M. & Duarte, J. H. The aryl hydrocarbon receptor: multitasking in the immune system. *Annu Rev Immunol* **32**, 403-432, doi:10.1146/annurev-immunol-032713-120245 (2014).

45 Qiu, J. *et al.* Group 3 innate lymphoid cells inhibit T-cell-mediated intestinal inflammation through aryl hydrocarbon receptor signaling and regulation of microflora. *Immunity* **39**, 386-399, doi:10.1016/j.immuni.2013.08.002 (2013).

46 Qiu, J. & Zhou, L. Aryl hydrocarbon receptor promotes RORgammat(+) group 3 ILCs and controls intestinal immunity and inflammation. *Seminars in immunopathology* **35**, 657-670, doi:10.1007/s00281-013-0393-5 (2013).

47 Fukumoto, S. *et al.* Identification of a probiotic bacteria-derived activator of the aryl hydrocarbon receptor that inhibits colitis. *Immunol Cell Biol* **92**, 460-465, doi:10.1038/icb.2014.2 (2014).

48 Moura-Alves, P. *et al.* AhR sensing of bacterial pigments regulates antibacterial defence. *Nature* **512**, 387-392, doi:10.1038/nature13684 (2014).

49 Daly, K. & Shirazi-Beechey, S. P. Microarray analysis of butyrate regulated genes in colonic epithelial cells. *DNA Cell Biol* **25**, 49-62, doi:10.1089/dna.2006.25.49 (2006).

50 Garrison, P. M., Rogers, J. M., Brackney, W. R. & Denison, M. S. Effects of histone deacetylase inhibitors on the Ah receptor gene promoter. *Arch Biochem Biophys* **374**, 161-171, doi:10.1006/abbi.1999.1620 (2000).

51 Gill, R. K. *et al.* Expression and membrane localization of MCT isoforms along the length of the human intestine. *Am J Physiol Cell Physiol* **289**, C846-852, doi:10.1152/ajpcell.00112.2005 (2005).

52 Hadjiagapiou, C., Schmidt, L., Dudeja, P. K., Layden, T. J. & Ramaswamy, K. Mechanism(s) of butyrate transport in Caco-2 cells: role of monocarboxylate transporter 1. *American journal of physiology. Gastrointestinal and liver physiology* **279**, G775-780 (2000).

53 Thibault, R. *et al.* Butyrate utilization by the colonic mucosa in inflammatory bowel diseases: a transport deficiency. *Inflammatory bowel diseases* **16**, 684-695, doi:10.1002/ibd.21108 (2010).

54 Flaveny, C. A., Murray, I. A., Chiaro, C. R. & Perdew, G. H. Ligand selectivity and gene regulation by the human aryl hydrocarbon receptor in transgenic mice. *Mol Pharmacol* **75**, 1412-1420, doi:10.1124/mol.109.054825 (2009).

55 Sokol, H. *et al.* Low counts of Faecalibacterium prausnitzii in colitis microbiota. *Inflammatory bowel diseases* **15**, 1183-1189 (2009).

56 Machiels, K. *et al.* A decrease of the butyrate-producing species Roseburia hominis and Faecalibacterium prausnitzii defines dysbiosis in patients with ulcerative colitis. *Gut* **63**, 1275-1283, doi:10.1136/gutjnl-2013-304833 (2014).

57 Kumari, R., Ahuja, V. & Paul, J. Fluctuations in butyrate-producing bacteria in ulcerative colitis patients of North India. *World J Gastroenterol* **19**, 3404-3414, doi:10.3748/wjg.v19.i22.3404 (2013).

58 Cushing, K., Alvarado, D. M. & Ciorba, M. A. Butyrate and Mucosal Inflammation: New Scientific Evidence Supports Clinical Observation. *Clin Transl Gastroenterol* **6**, e108, doi:10.1038/ctg.2015.34 (2015).

59 Alexeev, E. E. *et al.* Microbiota-Derived Indole Metabolites Promote Human and Murine Intestinal Homeostasis through Regulation of Interleukin-10 Receptor. *Am J Pathol*, doi:10.1016/j.ajpath.2018.01.011 (2018).

60 Zheng, L. *et al.* Microbial-Derived Butyrate Promotes Epithelial Barrier Function through IL-10 Receptor-Dependent Repression of Claudin-2. *Journal of immunology* **199**, 2976-2984, doi:10.4049/jimmunol.1700105 (2017).

61 Hungate, R. E. The anaerobic mesophilic cellulolytic bacteria. *Bacteriol Rev* **14**, 1-49 (1950).

62 Lakhdari, O. *et al.* Identification of NF-kappaB modulation capabilities within human intestinal commensal bacteria. *Journal of biomedicine & biotechnology* **2011**, 282356, doi:10.1155/2011/282356 (2011).

63 Bourriaud, C. *et al.* Lactate is mainly fermented to butyrate by human intestinal microfloras but inter-individual variation is evident. *J Appl Microbiol* **99**, 201-212, doi:10.1111/j.1365-2672.2005.02605.x (2005).

64 van Zundert, G. C. P. *et al.* The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J Mol Biol* **428**, 720-725, doi:10.1016/j.jmb.2015.09.014 (2016).

65 Wassenaar, T. *et al.* WeNMR: Structural Biology on the Grid. *Journal of Grid Computing* **10**, 743-767 (2012).

66 Wu, D., Potluri, N., Lu, J., Kim, Y. & Rastinejad, F. Structural integration in hypoxia-inducible factors. *Nature* **524**, 303-308, doi:10.1038/nature14883 (2015).

67 Kallberg, M. *et al.* Template-based protein structure modeling using the RaptorX web server. *Nat Protoc* **7**, 1511-1522, doi:10.1038/nprot.2012.085 (2012).

68 Ma, J., Peng, J., Wang, S. & Xu, J. A conditional neural fields model for protein threading. *Bioinformatics* **28**, i59-66, doi:10.1093/bioinformatics/bts213 (2012).

69 Chovancova, E. *et al.* CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. *PLoS Comput Biol* **8**, e1002708, doi:10.1371/journal.pcbi.1002708 (2012).

70 Ho, B. K. & Gruswitz, F. HOLLOW: generating accurate representations of channel and interior surfaces in molecular structures. *BMC Struct Biol* **8**, 49, doi:10.1186/1472-6807-8-49 (2008).

3.2. *Paper II: Identification of Bifidobacteria as novel activators of AhR pathway in human intestinal epithelial cells*

This second paper analyzes the butyrate-independent AhR activation induced by some bacterial strains. Within the commensal bacterial collection described in the first paper, some bacterial strains have been identified for activating AhR signaling in IECs, although they are not producers of butyrate. Considering the described indoles and indole-derivatives as main mediators for the bacterial-induced AhR activation, we selected the activatory strains among those described for not producing indoles. We thus evidenced that some of the tested Bifidobacteria, activated AhR signaling by a mechanism independent of both butyrate and indoles, thus describing a novel promising group of AhR activators. The Bifidobacteria are of particularly important for their described positive roles and their therapeutic application as probiotics, thus we wonder whether part of the beneficial effects could converge to the AhR activation. For this aim, preliminary experiments have been performed to decipher the signaling mechanism involved in the Bifidobacteria-induced activation of AhR and the final role on the epithelial barrier. Our preliminary results evidenced that some Bifidobacteria produced potential AhR ligand or ligand-precursors able to induce the expression of one readout gene for AhR activity (*CYP1A1*) and likely to stimulate the receptor translocation in the nucleus.

This project was funded by MetaCardis European project (FP7 - HEALTH-F4-2012-305312).

Identification of Bifidobacteria as novel activators of AhR pathway in human intestinal epithelial cells.

Ludovica Marinelli[1,2], Camille Martin-Gallausiaux[1,2], Fabienne Béguet-Crespel[1], Hervé M. Blottière[1, 3] and Nicolas Lapaque[1*].

1. Micalis Institute, INRA, AgroParisTech, Université Paris-Saclay, 78350 Jouy-en-Josas, France;

2. Sorbonne Universités, UPMC Univ Paris 06, IFD, 4 place Jussieu, 75252, Paris, cedex 05, France;

3. US 1367 MetaGenoPolis, INRA, Université Paris-Saclay, 78350 Jouy en Josas, France.

* Corresponding author:

Dr Nicolas Lapaque

INRA-MICALIS UMR1319, Bat 442, Domaine de Vilvert 78350, Jouy-en-Josas, France.

e-mail: nicolas.lapaque@inra.fr

Abstract

An increasing number of studies recently evidence the importance of the ligand activated transcription factor, aryl hydrocarbon receptor (AhR), as a critical regulator of immune and metabolic processes in the gastrointestinal track. In the gut, a main source of AhR ligands derives from commensal bacteria. However, many of the reported microbiota-derived ligands have been restricted to indole-derivatives. In this study, we aimed to identify new commensal bacterial strains that activates AhR pathway in human intestine. Using a screening with an AhR reporter system expressed in human intestinal epithelial cell line (IEC), we found that different strains belonging to Bifidobacteria activated AhR response. We showed that AhR ligand antagonists reduced the effects of Bifidobacteria on IEC suggesting that the activation could be driven by a produced AhR ligand not yet identified. Considering the known inability of Bifidobacteria in producing indoles, the main AhR ligands of microbial origins, our finding suggests that these bacteria are able to produce a yet unidentified AhR ligands. In conclusion, our findings suggest that (i) some Bifidobacteries activates AhR pathway and Cyp1a1 in human intestinal epithelial cell-lines (ii) some Bifidobacteries produce a potential novel ligand for AhR which might be an original mechanism for a potential beneficial effect of these bacteria on intestinal homeostasis.

Introduction

The mammalian gastrointestinal (GI) tract is colonised by a complex microbial community, referred as gut microbiota. The fine regulated host-microbiota interaction is particularly substantial for shaping host physiology and in particular mucosal barrier functions as well as the development and maintenance of the mucosal immune system (Spiljar et al. 2017; Wells et al. 2011; Cerf-Bensussan & Gaboriau-Routhiau 2010). Among the variety of commensal bacteria, Bifidobacteria have been extensively described in the host-bacteria crosstalk, in particular for their role in increasing epithelial cell integrity in vitro and in vivo (Hsieh 2015; Ewaschuk 2008; Lopez 2012). Bifidobacterium is a genus that dominates the intestine of healthy breast-fed infants, whereas in adulthood the levels are lower, but relatively stable. Numerous health-promoting effects have been ascribed to strains of the *Bifidobacterium* genus, among which the capacity to stimulate the immune system, in particular on the T-helper 1(TH1)/TH2 balance (Iwabuchi 2007), and the maintenance of epithelial cell integrity in vitro and in vivo (Hsieh 2015; Ewaschuk 2008; Lopez 2012). These roles are generally associated with the production of different metabolites, among which acetate (Fukuda 2011), conjugate linoleic acid (Raimondi 2016) and bacteriocins (Martinez 2013). Metabolites derived from commensal bacteria impact the host physiology by targeting surface receptors, such as G-protein-coupled receptors (GPRs) and intracellular receptors, such as different transcription factors, to induce signalling pathways that modify host gene expression and collectively impact on host metabolism and immune responses (S. Ranhotra 2017). Among the targeted intracellular receptor, the ligand-activated aryl hydrocarbon receptor (AhR) has recently gained considerable attention as a crucial modulator of mucosal immune and physical barrier. AhR is a member of the basic helix-loop-helix PER-ARNT-SIM (bHLH-PAS) family, initially identified as the hepatic intracellular protein that bounds with high affinity to the environmental halogenated contaminant 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). However, it became evident that AhR is able to bind a huge variety of ligand, including the microbial-derived indoles (Hankinson 1995; Barouki et al. 2007; Barouki et al. 2012; Hubbard, Murray, Bisson, et al. 2015)(Jin et al. 2014; Hubbard, Murray & Perdew 2015; Murray & Perdew 2017). Upon ligand-binding, the cytoplasmic AhR induces the expression of genes with promoters containing a xenobiotic-response element (XRE) consensus sequence. Genes such as cytochrome P450 family 1A1 (CYP1A1) and the repressor AhRR are regulated by AhR activation (Hankinson 1995; Fujii-Kuriyama et al. 1994; Stockinger et al. 2014). AhR activation has been largely described to be implicated in colonic stem cells proliferation, epithelial barrier functions, maintenance of intraepithelial lymphocytes (IEL), innate lymphoid cells (ILC) and FOXP3 regulatory T cells (Treg) (Stockinger et al. 2014; Zelante et al. 2016; Monteleone et al. 2013; Lee et al. 2012; Gagliani et al. 2015; Shimada et al. 2013; Bansal et al. 2010). Interestingly, AhR signalling and known AhR ligands are low in inflammatory bowel disease (IBD) patients, highlighting the clinical relevance of the AhR pathway in this disease (Lamas et al. 2016; Monteleone, Rizzo, Sarra, Sica, Sileri, Biancone, MacDonald, Pallone & Monteleone 2011). Recently it has been described the role of AhR in enhancing IECs barrier functions via the increase of the IL10

receptor expression and the enhancement of tight junctions integrity through the regulation of Notch1 (Liu et al. 2018; Yu et al. 2018). These results highlighted the importance of AhR ligands in modulating host gut immune homeostasis and prompted us to identify new activators of the AhR activating pathway in IECs, other than the described indoles.

In this study, we aimed to identify new commensal bacterial strains and metabolites that influence AhR-dependent genes activation, using a screening with an AhR reporter system expressed in the human intestinal epithelial cell line HT-29. We thus tested the bacterial supernatants from a collection of Bifidobacteria on an AhR reporter system and found that some Bifidobacteria activate AhR signalling. Antagonists, blocking the binding of AhR ligand, impaired the Bifidobacteria-induced activation of AhR reporter system and the up-regulation of *CYP1A1* gene expression, highlighting that this group of bacteria is likely to produce a potential AhR ligand, able to activate the AhR signalling in IECs. We thus identify for the first time that Bifidobacteria are a promising group of bacteria activating AhR signalling in IECs by a mechanism not yet completed decrypted.

Results

Some Bifidobacteria spp. enhance AhR activity

Bifidobacteria are commensal bacteria, residing in the intestine in close contact with intestinal epithelia. Assuming that AhR activators are secreted in the intestinal lumen (Jin 2014), we screened on the intestinal epithelial cell line HT29 stably expressing an AhR reporter system, a collection of bacterial supernatants from commensal Bifidobacteria. The collection included 16 strains belonging to 12 *Bifidobacterium* species and 3 distinct phylogenetic groups (*Bifidobacterium adolescentis* group, *Bifidobacterium pseudolongum* group and *Bifidobacterium breve* & *Bifidobacterium longum* group, accordingly to Felis and Dellaglio 2007) and one species not assigned to any group (*Bifidobacterium bifidum*) (Supplementary Figure S1, Supplementary Table 1). We evidenced that the non-inoculated bacteria media (M58, M104, LYBHI1 and LYBHI4JR) activated the AhR reporter system and we confirmed this results in FBS-free condition, to avoid additional bias induced by potential AhR ligands contained in the FBS-supplemented cell culture media (Supplementary Figure S2). The rich composition of the bacterial media (Supplementary Table 2) is likely to be at the origin of the basal AhR activation and probably due to some AhR ligands already present in the bacterial media. M58 and M104 showed an AhR activation closed to the one observed with the prototypical ligand TCDD. However the two bacterial media, LYBHI1 and LYBHI4JR, showed a far higher activation of the reporter system. As it is conceivable that the strong background activation of AhR induced by LYBHI4JR and LYBHI1 (Supplementary Figure S2) could mask the bacterial activity, we choose to exclude the bacterial culture performed on these media from the data sample. Additionally, the presence of carbohydrates (maltose, starch and cellobiose) in LYBHI4JR and LYBHI1 in contrast with the more proteic composition of M58 (casein peptone, meat extracts) (Supplementary Table 2), stimulate a different utilisation of the available substrates with likely consequence on basal and bacterial-induced AhR activations (Figure 1). By analysing only the bacterial cultures performed in M58 and M104 we did not evidenced differences in AhR activation comparing the different phylogenetic groups. However, apart from media composition and phylogenetic groups, differences in the induced AhR activity emerged between the two strains of *Bifidobacterium breve* (ATCC15701 and DM 20091, both in M104) (Figure1), but the origin for this discrepancy remains an open question. Beside *B. breve* DSM20091, *B. ruminantium* and *B. animalis subsp. animalis* have been also identified as non-activators in the reporter system (Figure 1).

Some Bifidobacteria activate the AhR-dependent gene CYP1A1 in intestinal epithelial cell lines

The activation of the AhR pathway induces the expression of different AhR-dependent genes, including *CYP1A1*. To confirm that some Bifidobacteria activate AhR pathway, we selected *B. choerinum*, *B. longum* subsp. *longum* and *B. bifidium* and *B. breve* ATCC15701 to evaluate the *CYP1A1* mRNA expression induced by these strong activators. With the sole exception of *B. breve* ATCC15701, the

three tested strains induced *CYP1A1* mRNA expression in HT29 cells, confirming the role on AhR pathway activation.

The bacterial-induced activation is independent of acetate and lactate productions

Bifidobateria have been largely described as producers of acetate and lactate (Pokusaeva 2011). We thus hypothesized that acetate or lactate concentrations in the supernatants of these bacteria could explain the activation of the AhR pathway. We therefore quantified the concentrations of both metabolites in some bacterial supernatants and we confirmed the ability of these strains to produce acetate in our culture conditions (Supplementary Table 1). Consequently, we tested pure acetic acid and lactic acid on HT29-AhR reporter cell line at concentrations ranging from 0.125mM and 8mM, coherent with those detected in the supernatants (Supplementary Table 1, Figure 2). We observed no impact of both metabolites on our reporter system (Figure 2), suggesting that the bacterial-induced activation of AhR is independent of acetate and lactate production.

Some Bifidobacteria produce potential novel AhR ligands

As AhR is a ligand-activated transcription factors, we assessed if Bifidobacteria could activate AhR signalling by the production of a metabolites acting as AhR ligand. We thus tested the Bifidobacteria-induced AhR response in presence of two well-described antagonists for the AhR-ligand binding (CH-223191 and GNF-351) in HT29-AhR cell line (Zhao 2010; Smith 2011). Our result shows that the two inhibitors decreased the activation of AhR induced by all the tested bacterial strains in HT29-AhR reporter cell lines (Figure 4). In particular, AhR activity induced by *B. choerinum* was halved by both inhibitors while in the Bifidobacteria, activators of AhR, GNF-351 showed a higher inhibitory potential. The inhibition observed incubating CH-223191 and GNF-351 with M58 and M104, suggests the presence of probable ligands already present in both media. Although, the stronger activation especially induced by *B. choerinum* and *B. bifidum* and the observed inhibition induced by both antagonists, could not be completely explained by the basal inhibition of the culture media. Consequently, we suspect that other possible ligands, different from those present in the media, are responsible for the far higher activation induced by the bacteria. Alternatively, as proposed for other AhR activators, weak ligands or ligand precursors could be present in the culture media and converted in stronger agonists by the microbial metabolization. However, further investigation are needed to better discriminate from the basal activation of the non-inoculated media, to the bacterial cultures.

It is well established that ligand binding triggers the accessibility of the nuclear localisation signal on the AhR N-terminus that consequently initiates the AhR nuclear translocation event (Henry 2003). Thus, we investigated, by nuclear translocation assay, the human AhR agonist potential of Bifidobacteria-derived metabolites. Sub-cellular localisation of AhR in HT-29 incubated with bacterial supernatants at

1h (Figure 6), 3h and 6h (not shown) was assessed by immunoblotting assay. Our results showed an accumulation of AhR proteins in the nucleus upon treatment with bacterial supernatant during 1h, consistent with the action of an AhR agonist (Figure 6).

Altogether, our results suggested that the activation of the AhR reporter system by some Bifidobacteria is likely due to the production of active metabolites acting as AhR ligands, likely to induce the AhR translocation to the nucleus.

Analysis of the Bifidobacteria for prediction of tryptophanase expression

Despite the huge literature on AhR ligands, the described AhR ligands of bacterial origins are limited. Among them, indoles are produced by tryptophanase (TnaA), which can reversibly convert tryptophan into indole, pyruvate, and ammonia (Lee 2010). Many Gram-positive and Gram-negative bacteria encode a single copy of the *tnaA* gene in their chromosome and produce indole. Although most organisms contain the tryptophan biosynthesis pathway or express a TnaA homologue, to date, only those encoding *tnaA* can synthesize indole. To identify if the selected Bifidobacteria are equipped for the production of indoles through tnaA, we analysed the predicted tryptophanase function (4.1.99.1) in publicly available genomes of selected species, using the Integrated Microbial Genomes (IMG) system. Due to the unavailability of the genomes for the exact tested strains, the analysis was performed at the species level selecting genomes (preferring the finished one when possible) for the same species among the complete available genomes (Table 2). The indole-producer *Clostridium indolis* has been used as positive controls for annotated tryptophanase (Table 2).

Accordingly to this analysis, selected Bifidobacteria species are not annotated for expressing tnaA (4.1.99.1) (Table 2). These results were further confirmed with the annotated pathways and genomes in KEGG and crosschecked with bibliographic evidences (Supplementary Table 3). Coherently, the activation induced by Bifidobacteria is likely due to the production of active metabolites different from the described indoles.

Discussion

The epithelial cells of the gastrointestinal (GI) tract is at the interface between the host and the microbial community, in a first line for establishing a human-microbes crosstalk able to modulate intestinal homeostasis. A ligand-activated transcription factor called aryl hydrocarbon receptor (AhR), ubiquitously expressed in eukaryotes, has recently gained rising consideration regarding intestinal homeostasis. AhR activation regulates both adaptive immunity and mucosal barrier functions (Quintana & Sherr 2013; Stockinger et al. 2014) and emerging evidences highlight that AhR directly impacts intestinal epithelial cells (IECs) leading to increased barrier functions and the regulation of excess AhR ligands in the gut (Schiering et al. 2018). The roles of AhR are generally mediated by ligand binding and the GI tract constitute a rich source of these molecules, derived from food, bacteria or endogenously produced.

Despite the huge literature on the structurally pleiotropic nature of AhR ligands, only few are bacterial-derived molecules have been identified. Amongst them, indoles and other tryptophan derivatives are produced by a variety of bacteria including some Lactobacilli (Firmicutes) and Proteobacteria (Zelante 2013). Apart of indoles, the probiotic derived molecule DHNA (1.4-Dihydroxy-2-naphthoic acid), produced by *Propionibacterium freudenreichii*, has been shown to induce anti-microbial proteins RegIIIβ and RegIIIγ through the AhR activation in the intestine of mice, with a consequent alteration of the intestinal microbial community and inhibition of experimental-induced colitis (Fukumoto et al. 2014). Given the abundance and extensive metabolic capacity of gut microbiome, it is likely that other active metabolites are present in the GI tract and could potentially stimulate AhR pathway.

Bifidobacteria is a dominant genus in the intestine of healthy breast-fed infants and also stably present in adults (Biavatti 2006, Voreades 2014). Numerous health-promoting effects have been ascribed to strains of the *Bifidobacterium* genus thus justifying their large use as probiotics in intervention strategies to address many ill conditions (Menard 2008, Di Gioia 2014, Biagi 2012). Among the positive role of these bacteria, the stimulation of systemic and intestinal immunity (Menard 2008) has been evidenced together with a positive role in pathologies such as allergies, celiac disease, obesity, diarrheas, infections or necrotizing enterocolitis (Di Gioia 2014). Considering the overall positive role of Bifidobacteria in promoting health, we address the question if these bacteria are be able to stimulate AhR pathway in IECs that could explain parts of their described positive effects. By screening a collection of supernatants from Bifidobacteria species, we identified as activators of AhR signalling in HT29 cells *B. dentium DSM20054, B. catenulatum, B. dentium ATCC25734, B. angulatum, B. adolescentis, B. pseudocatenulatum, B. breve ATCC15701,B. longum subsp. infantis, B. longum subsp. longum, B. choerinum and B. bifidum* (Figure 1). Additionally we evidenced that *B. choerinum, B.bifidum* and partially *B. longum subsp. longum* are able to induce the expression of the AhR-dependent gene *CYP1A1*, further confirming the activation of the AhR signalling pathway.

Surprisingly, differences in the induced AhR activity emerged between the two strains of Bifidobacterium breve (ATCC15701 and DSM20091 both in M104) (Figure 1). Unfortunately, of the current recognized bifidobacterial species, only few have been entirely sequenced (Barrangou 2009; Kim 2009; Lee 2008; Schell 2002; Sela 2008; Ventura 2009; Turroni 2010; Hao 2011; Zhurina 2011) and, concerning the B. breve, only the reference genome (DSM20213) have been recently annotated (NCBI Reference Sequence: NZ_AP012324.1). Consequently, the comparison between the ATCC15701 and DSM20091 genomes was not possible and the source of this evidenced difference remains an open question.

An important function of the *Bifidobacterium* genus that contributes to gut homeostasis and host health is the production of acetate and lactate during carbohydrate fermentation, in turn converted into butyrate by other colon bacteria through cross-feeding interactions (De Vuyst 2011 and 2014; Rivière 2015). Consequently, we wonder if these acetate and lactate could be at the origin of the Bifidobacteria-induced activation of AhR. Firstly, we quantified the produced concentrations of acetic acid and lactic acid in bacteria supernatants confirming their production in our experimental conditions, then we tested the pure organic acid on AhR reporter system. By luciferase assay we observed that both acetate and lactate were not able to induce AhR activation in HT29-AhR cell line, suggesting that the involved mechanism should imply other bacterial-produced metabolites.

Excluding the hypothesis that acetate and lactate could activate AhR, we wonder if Bifidobacteria could produce an AhR ligand directly activating the signalling pathway. Numerous studies evidenced that the predominant biological activities of the AhR are through ligand binding and despite the AhR association with dioxins, structurally diverse metabolites from the diet, bacteria or produced by the host have been reported as capable of binding to human AhR (Hubbard, Murray, Bisson, et al. 2015; Flaveny et al. 2009; Murray & Perdew 2017). To assess if Bifidobacteria activate AhR by producing ligands, we incubated two well-describe specific antagonists (CH-223191 and GNF-351) in presence and absence of the bacterial supernatants. Our results suggested that some strains of Bifidobacteria (mainly *B. bifidum*, *B. choerinum*, *B. angulatum* and *B. dentium*) produce potential AhR ligands as the inhibition of ligand binding by CH223191 and GNF-351, resulted in a reduction of the bacterial induced AhR response. Additionally, our results showed that *B. bifidum* and *B. choerinum*, together with *B. bifidum* stimulates the AhR translocation in the nucleus, consistent with the role of other described AhR agonists.

Although not the entire bacterial collection has been investigated for *CYP1A1* mRNA expression and AhR nuclear translocation, our preliminary results on some Bifidobacteria strains belonging to different phylogenetic groups (mainly *B. bifidum*, *B. choerinum*, *B. angulatum* and *B. dentium*), provide promising evidences for a novel role of Bifidobacteria in activating AhR signalling in IECs.

The depletion in AhR ligands indoles from a low-tryptophan diet, results in a higher susceptibility of mice to experimentally induced colitis compared to mice fed a normal tryptophan diet (Hashimoto et al., 2012). Accordingly, indole-3-proprionic acid, an AhR ligand, was found selectively diminished in

circulating serum from human subject with active colitis (Alexeev et al. 2018). Additionally, AhR activation by microbial derived indole 3-aldehyde, stimulates the IL-22 and antimicrobial production in the gut with a concomitant protection against *Candida albicans* infection in mice (Zelante 2013). Interestingly, some described positive role of Bifidobacteria in intestinal homeostasis exhibit overlapping activities with AhR ligands. Indeed, a decrease in the relative abundances of *Bifidobacterium* species in human colon has been associated with antibiotic-associated diarrhea, IBD, obesity, allergies, and regressive autism (Di Gioia 2014; Grimm 2014) and the depletion in acetate-producing Bifidobacteria was described to increase the susceptibility to the enteropathogenic infections and promote excessive intestinal inflammation in mice (Fukuda 2011, Gao 2015). Hence, it is possible that some of the positive effects of Bifidobacteria in human gut are mediated by AhR. Coherently, a recent report described an increased expression of *AhR* gene in CD4+ T cells co-culture with *B. breve*-treated intestinal dendritic cells further propounding for a role of Bifidobacteria on AhR signalling, although the exact mechanism involving AhR has not been investigated (Jeon 2012).

Some evidenced suggested that the positive roles of Bifidobacteria seems not characteristic of the entire genus or certain species, but are rather strain-specific. Interestingly, in our screening most of the tested bacteria induced the activation of AhR at least in reporter system, with no discrimination among phylogenetic groups, suggesting that a metabolite more generally produced within the genus could be involved in the activation.

In conclusion, we show that Bifidobacteria are a novel group of bacteria activating AhR pathway, stimulating CYP1A1 expression through the production of yet unidentified metabolites that likely bind to AhR. Thus, our preliminary results provide encouraging evidenced for the production of a potential AhR ligand by some member of Bifidobacteria genus (*B. bifidum*, *B. choerinum*, *B. angulatum* and *B. dentium*) which is, to our knowledge, an original role for Bifidobacteria.

Materials and Methods

Cell Culture of human colonic cell lines

The human epithelial cell lines HT-29 were obtained from the American Type Culture Collection (ATCC, Rockville, MD), grown in RPMI 1640 GlutaMAX™ supplemented with 10% of heat-inactivated fetal bovine serum (FBS, Lonza). RPMI was supplemented with 50 IU/mL penicillin, 50 µg/mL streptomycin and 10%, 100mM Hepes, 10mM nonessential amino acids. HT-29 were grown at 37°C in a humidified 5% CO2 atmosphere. All culture media and supplements were supplied by Gibco (ThermoFisher). Mycoplasma contamination was regularly tested using MycoAlert (Lonza) and PlasmoTest (Invivogen).

Production of Stable AhR-luciferase Reporter Cell-Lines

pGL4.43[luc2P/XRE/Hygro] (Promega) was used to establish HT29-AhR and Caco2-AhR reporter cell-lines by electroporation using the Nucleofector® device (Lonza) according to the manufacturer's recommendations. Stable AhR reporter cell lines were selected using Hygromycin (600 µg/ml for HT29 cell line, InvivoGen) and validated using TCDD at 10nM final concentration.

Culture of commensal bacteria, preparation of supernatants and SCFA concentration assessment

16 Bifidobacterium spp. strains belonging to 12 Bifidobacterium species from the in-house INRA-Micalis collection or from DSMZ were grown. Anaerobic culture conditions were done accordingly to the Hungate method (Hungate 1950). Screened strains, corresponding growth media, optical densities (OD600), organic acid concentrations are listed in Supplementary Table 1 and composition of home-made growth media is listed in Supplementary Table 2. Bacterial cultures were cultured to reach the maximum OD. Bacterial supernatants were harvest after centrifugation at 5,000 × g for 10 min and filtered on a 0.22µm PES filters and stored at -80°C. Quality controls were performed using Gram staining method, aerobic growth test and fresh observation on microscope. Non-inoculated bacteria culture medium served as control. Concentrations of organic acids produced by cultured bacteria were measured by HPLC and gas chromatography as described by Lakhdari (2011).

Luciferase Reporter and Cell Viability Assays

For the bacterial screening, HT-29-AhR cells were seeded at 3×10^4 cells per well in white 96-well plates (Corning). After 24 h from seeding, cells were stimulated during 24 hours with 10 µL of bacterial supernatant or un-inoculated media in a total culture-volume of 100 µL per well (10% vol/vol). The screening was performed in triplicates and for almost all the samples, experiments were performed at least with two biological replicates. Additionally, when possible, some strains were grown in different bacterial media. All the experiments performed after the main screening (confirmation of active bacteria, inhibitors, RNA and protein extractions) were performed in non-FBS-supplemented media. To do that, 24h after seeding the culture media was replaced with a non-FBS-supplemented RPMI. The cells were

then stimulated with 10µL of reagents or bacterial supernatants diluted in non-FBS conditions in a total culture-volume of 100 µL per well (10% vol/vol). Follow-up experiments were performed in triplicates. Luciferase activity was quantified as relative luminescence units (RLU) using a microplate reader (infinite® 200 plate reader, TECAN) and the Neolite™ (PerkinElmer) Luciferase Assay System according to the manufacturer's instructions. The AhR activation was normalised on non-inoculated bacterial media or untreated/vehicle-treated cells for bacterial supernatants as indicated in the figures. The results were expressed as luciferase fold change. Cell viability was assessed by MTS measurement using the CellTiter 96 Aqueous One solution (Promega) according to the manufacturer's recommendations.

Reagents

All agonists, antagonists and drugs tested were dissolved in a proper vehicle (DMSO, glycerol, water, PBS or ethanol) following the manufacturer's recommendations. The final concentration used for vehicles had not detectable effect on metabolic activity of the cells. Acetic acid and Lactic acid were from Sigma and used in a range of concentration from 0.125 to 8mM. AhR agonist: 2,3,7,8-Tetrachlorodibenzodioxin (TCDD 10nM, Sigma). AhR antagonists: CH-22319 (1µM, Millipore/Calbiochem), GNF-351 (1µM, Millipore/Calbiochem).

Real-Time PCR

Cell lines were seeded in 12-well culture plates at densities of 0.5×10^6 cells per well. The cells were seeded in FBS-supplemented media then, after 24h, the media was replaced with a non-FBS-supplemented and cells incubated during 24h before stimulation. After stimulation time of 6h, total RNA was extracted using RNeasy mini-Kit (Qiagen) according to manufacturer's recommendations. cDNA was synthesized from 2 µg of RNA using the High-Capacity cDNA Archive Kit (Applied Biosystems). qPCRs were carried out using an StepOne (Applied Biosystems) thermal cycler in a reaction volume of 20 µL with Taqman gene expression assay probes: CYP1A1: Hs01054796_g1; β-Actin: Hs99999903_m1. CYP1A1 expression relative to control expression was determined by the $2^{-\Delta\Delta Ct}$ method using β-actin as control gene. Experiments were performed in technical triplicates.

Cytoplasmic and nuclear protein extraction

HT-29 cells were seeded at densities of 0.5×10^6 cells per well in 12-well-plates. 24h after seeding the media was replaced with a non-FBS-supplemented RPMI and cells incubated during 24h prior to stimulation. For compartments separation, nuclear and cytoplasmic extracts were prepared with NE-PER Nuclear and Cytoplasmic Extraction Reagent Kit (ThermoFisher) according to the manufacture instructions. CER I and NER buffers were supplemented with protease inhibitor cocktail (cOmplete™ ULTRA Tablets, Mini, EASYpack Protease Inhibitor Cocktail, Sigma) prior to use.

Western Blot analysis

Protein extracts were run in 10% SDS-PAGE gels and transferred onto PVDF membranes by liquid transfer (Transfer buffer: 192mM Glycine, 25mM TrisBase, 20% methanol) at 200mA during 90minutes. Membranes were blocked overnight in TBST+4% BSA (Sigma). Primary antibodies were incubated overnight at 4°C: anti-AhR (1:500, mouse mAb, clone RTP1, ThermoFisher), anti-Lamin A/C (1:2000, mouse mAb, Cell Signaling), anti-GAPDH (1:4000, mouse mAb, Santa Cruz), anti-Lamp1 (1:2000, mouse mAb, H4A3 from the Developmental Studies Hybridoma Bank (DSHB), H4A3 was deposited to the DSHB by August, J.T./Hildreth J.E.K. (DSHB hybridoma product H4A3). Secondary mouse horseradish peroxidase-coupled antibody (DAKO) was successively incubated at room temperature for 2h before detection with the Clarity Western ECL Substrate using the Chemidoc MP System (Bio-Rad). Quantifications were performed using the image Lab software (Bio-Rad). AhR nuclear protein levels were normalised to Lamin A/C protein levels. Lamp1 and GAPDH were used as purification controls for the cytoplasmic proteins.

Comparative analysis

Comparative analysis of genomes vs. functions were performed using Integrated Microbial Genomes (IMG) system and KEGG Pathway database.

Figure legends

Figure 1: Screening of in-house collection of Bifidobacteria on HT29-AhR cell line. HT29-AhR reporter cells were incubated with bacterial supernatants or relative non-inoculated bacterial media for 24h (10% vol/vol) in non-FBS supplemented RPMI. AhR activation was measured by luciferase activity and expressed as fold increase (± SD from triplicates) toward its control (non-inoculated bacterial media), sorted by phylogenetic groups (*B. adolescentis* group, *B. breve* and B. longum group), *B. pseudolongum* group, NA: Not Assigned group). Colours refer to bacterial media: M58 in green, M104 in orange, LYBHI1 in pink, LYBHI4JR in light blue.

Figure 2: Bifidobactera activate the expression of the AhR-regulated gene *CYP1A1*. HT-29 cells were treated with bacterial supernatants or non-inoculated culture media during 6h. The mRNA expression of *CYP1A1* was determined by qRT-PCR. *CYP1A1* relative expression to control is determined by the $2^{-\Delta\Delta Ct}$ method using *β-ACTIN* for normalisation. Data are expressed as means ± SEM of at least three distinct experiments, performed in triplicate.

Figure 3: Bifidobacteria activate the AhR pathway independently of acetate and lactate production. HT29-AhR reporter cells were incubated with a range of concentration of acetate and lactate rising

(0.125mM to 8mM). Data are expressed as luciferase fold (± SEM) of at least three independent experiments, normalised on untreated cells.

Figure 4: AhR antagonists inhibit AhR activation by Bifidobacteria. HT29-AhR reporter cells were incubated with AhR ligand antagonists (CH-223191 1µM and GNF-351 1µM) prior to stimulation with bacterial supernatants for 24h. Activation of the AhR reporter system was measured by luciferase activity, normalised on non-incubated culture media and expressed as fold increase means (± SEM) of at experiments performed in technical triplicates.

Figure 5: Bifidobacteria stimulate AhR nuclear translocation. HT-29 cells were incubated with bacterial supernatants during 1h. Nuclear extracts were assessed for AhR protein expression. Lamin A/C was used as control for nuclear extract preparation (left panel). Relative quantification of AhR protein level in nuclear extracts is expressed as fold-change (± SEM) to un-stimulated cells (control) after normalisation on Lamin A/C level.

Table legends

Table 1: Comparative analysis of some available genomes for tested Bifidobacteria species and *Clostridium indolis*, regarding the expression of tryptophanase (EC: 4.1.99.1) using IMG. The numbers in the tryptophanase column indicates the presence (1) or the absence (0) in the selected genomes. Sequencing status, Sequencing center, IMG Genome ID, Genome Size and Gene Count assembled are displayed as provided by IMG.

Table 2: Cross-check between KEGG annotation and bibliographic references for the prediction of tryptophanase activity in the tested Bifidobacteria strains.

Supplementary Figures

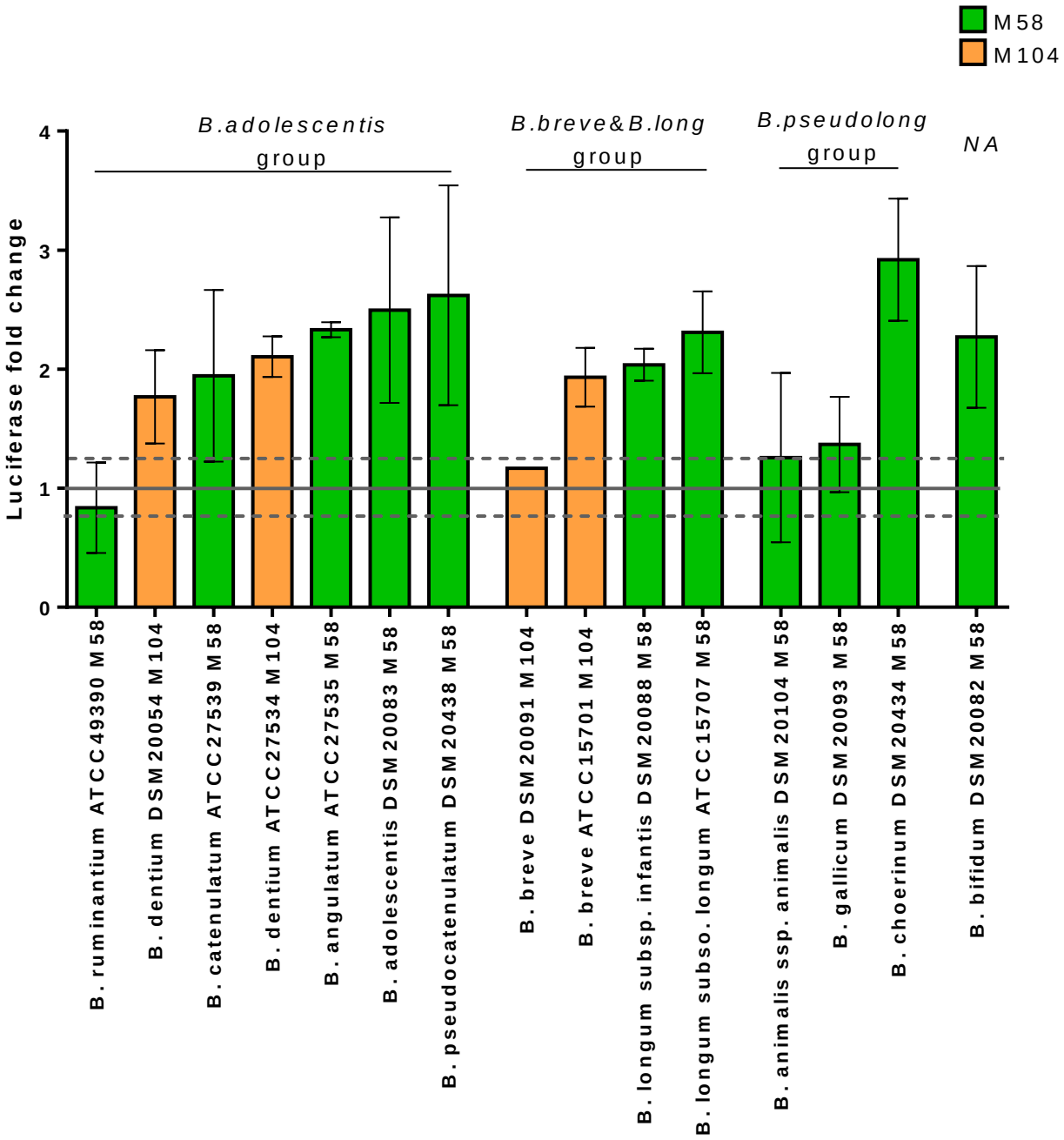
Figure S1: Screening of in-house strain collection of Bifidobacteria on HT29-AhR cell line in presence of FBS. AhR activation is expressed as the fold increase (± SD) toward its control (non-inoculated bacterial media), sorted by bacterial strains and phylogenetic groups (*B. adolescentis* in purple, *B. breve* and *B. longum* in light blue, *B. pseudolongum* group in green, NA: Not Assigned group in gray). HT29-AhR cells were exposed to bacterial supernatants or relative non-inoculated bacterial media for 24h (10%vol/vol).

Figure S2: Non-inoculated culture media activated AhR on HT29-AhR cell line, in absence of FBS. AhR activation is expressed as the fold increase (± SD) toward RPMI-stimulated cells and compared to the positive control TCDD (10nM). HT29-AhR cells were exposed to bacterial media, TCDD or RPMI for 24h (10%vol/vol).

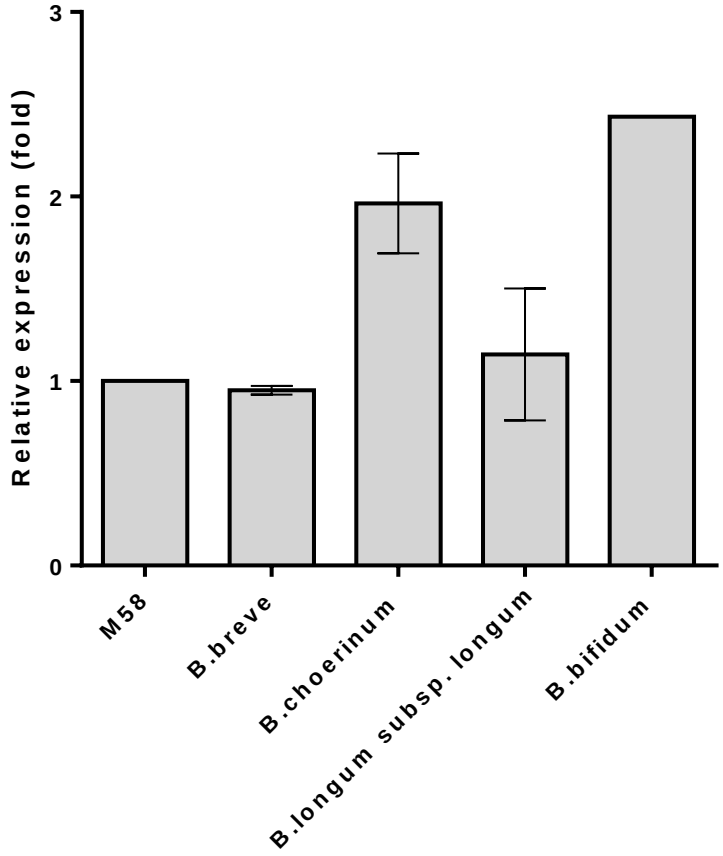
Supplementary Tables

Supplementary Table S1: List of all tested bacterial strains, bacterial growth (measured by optical density 600, OD600), pH of conditioned medium, organic acid concentrations and AhR activity (as fold increase as compared to control medium)

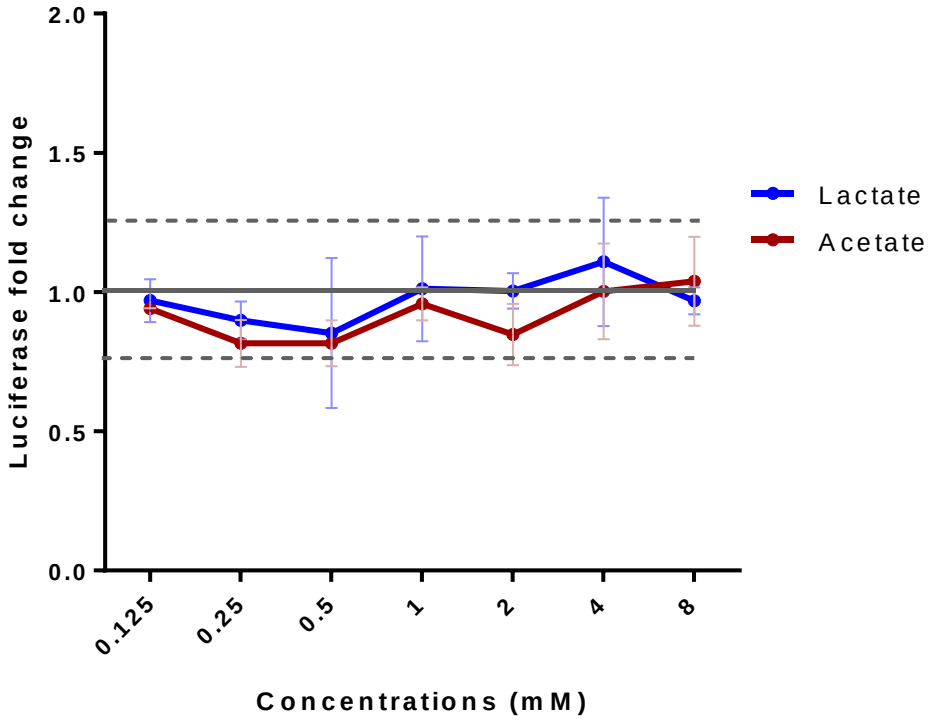
Supplementary Table S2: Composition of bacterial media used in the study.



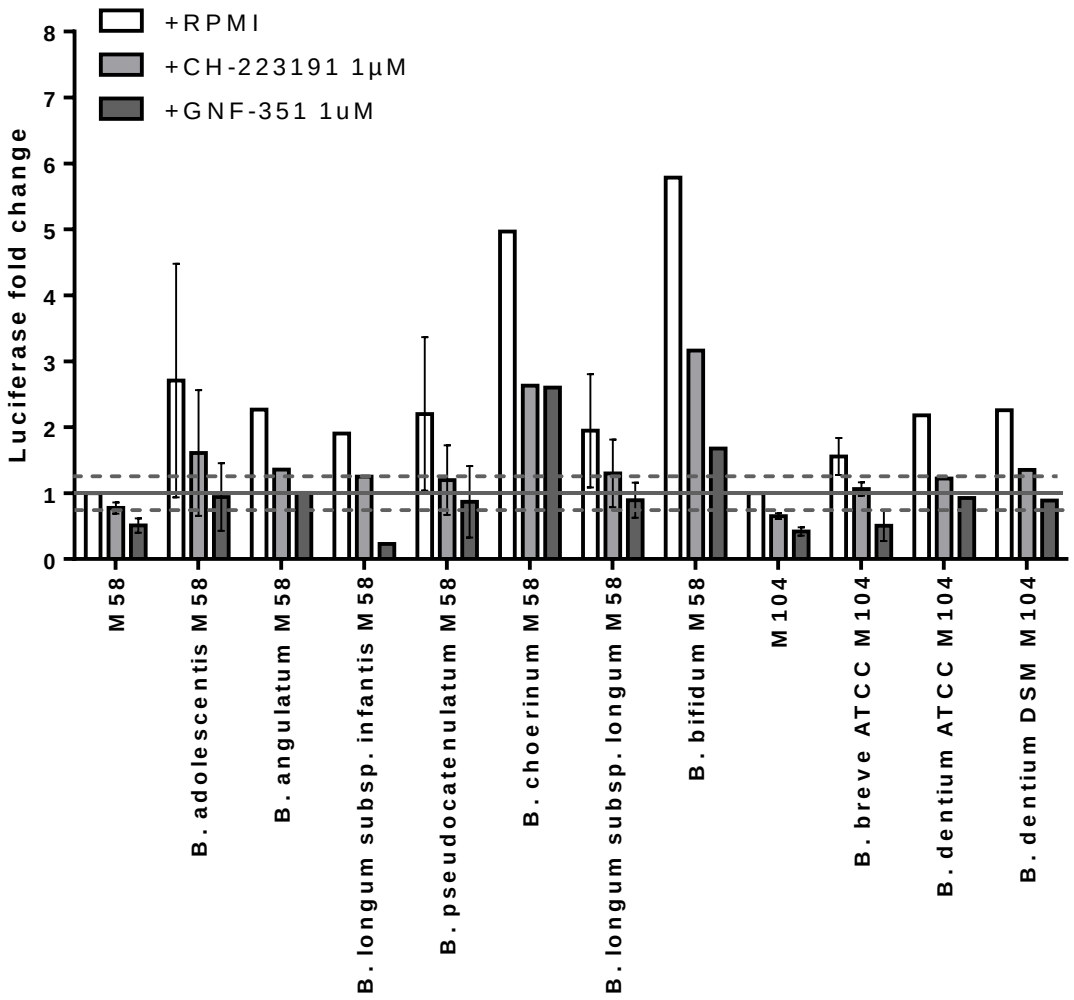
Marinelli *et al.*
Figure 1



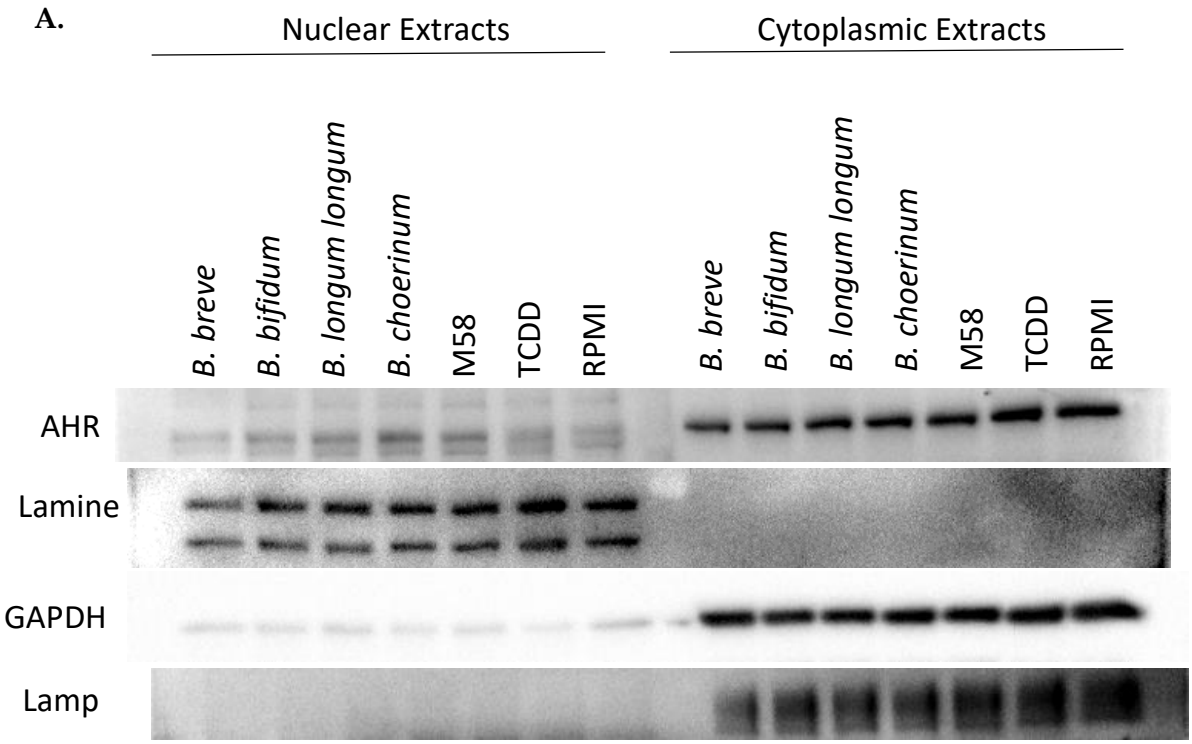
Marinelli *et al.*
Figure 2

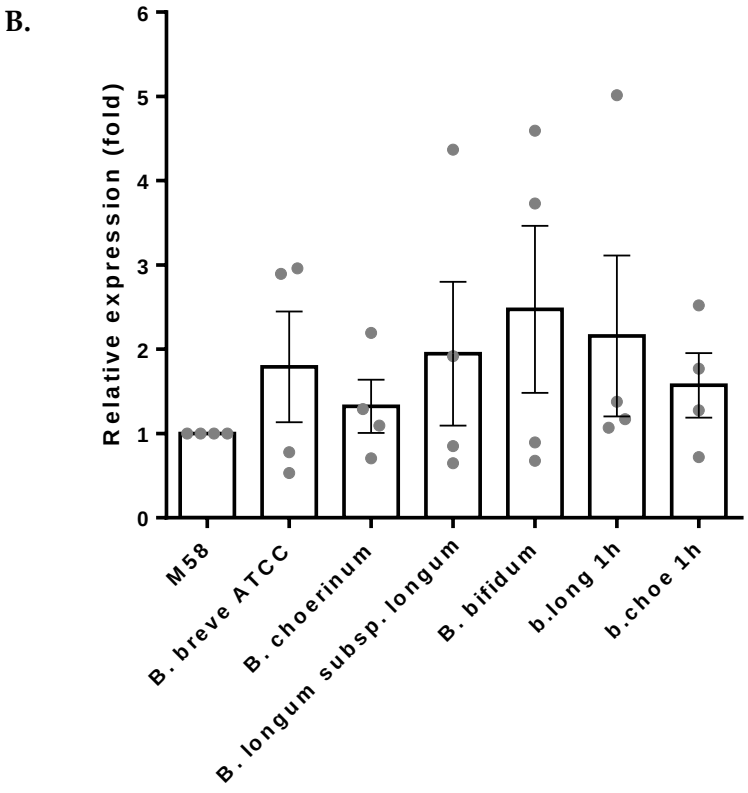


Marinelli *et al.*
Figure 3



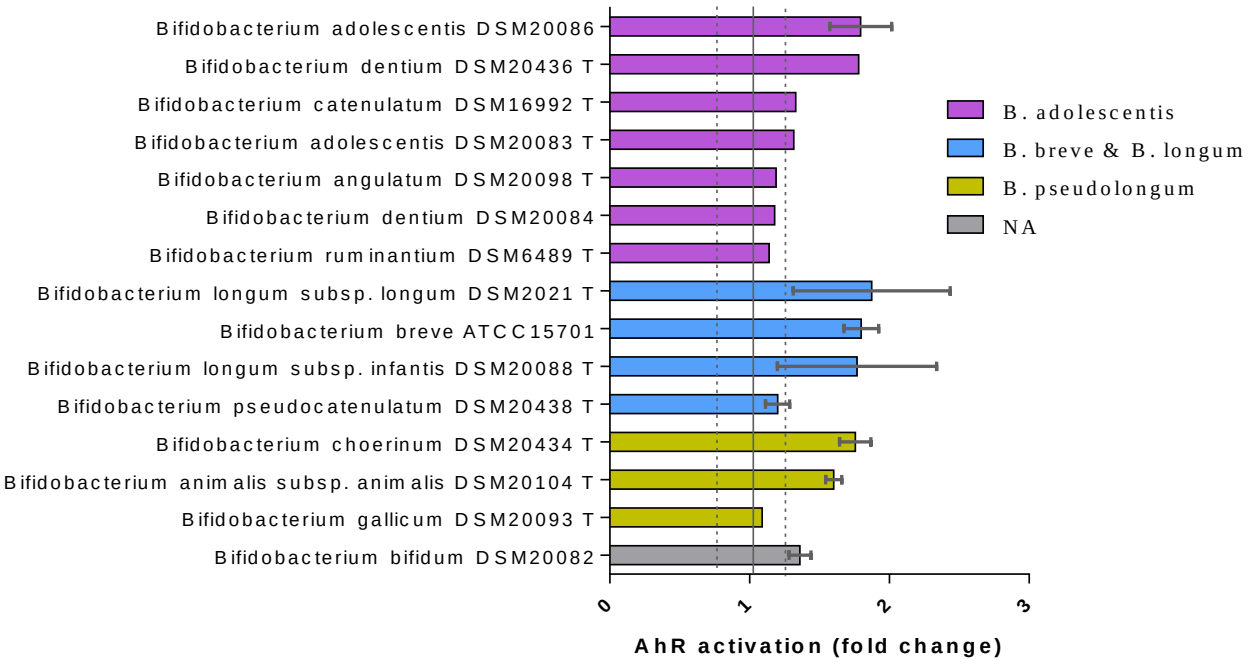
Marinelli *et al.*
Figure 4



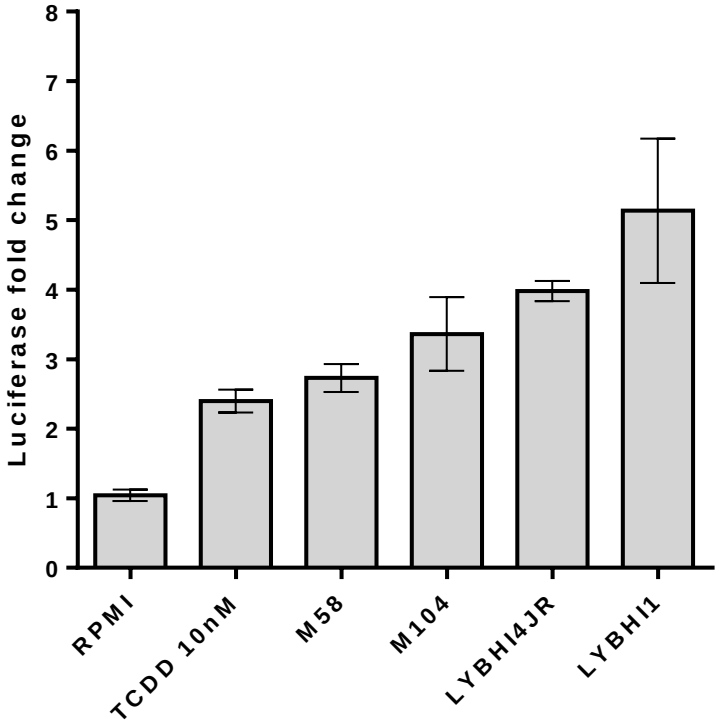


.

Marinelli *et al.*
Figure 5



Marinelli *et al.*
Supplementary Figure 1



Marinelli *et al.*
Supplementary Figure 2

Marinelli *et al.* Table 1

Genome Name / Sample Name	Sequencing Status	Sequencing Center	IMG Genome ID	Genome Size	* assemble	Gene Count	* assemble	EC:4.1.99.
Clostridium indolis DSM 755	Permanent Draft	DOE Joint Genome Institute (JGI)	2524614838	6383701	5902	1		
Bifidobacterium adolescentis 22L	Finished	University of Parma	2597490324	2203222	1817	0		
Bifidobacterium adolescentis ATCC 15703	Finished	Gifu University	639633010	2089645	1709	0		
Bifidobacterium adolescentis BBMN23	Finished	China Agricultural University	2648501305	2173720	1831	0		
Bifidobacterium angulatum DSM 20098	Permanent Draft	Inner Mongolia Agricultural University	2700989373	1990282	1635	0		
Bifidobacterium angulatum DSM 20098	Permanent Draft	Washington University in St. Louis	2562617143	2007108	1810	0		
Bifidobacterium angulatum GT102	Finished	Vavilov Institute of General Genetics	2645727667	2044762	1678	0		
Bifidobacterium animalis animalis ATCC 25527	Finished	Pennsylvania State University	2513237236	1932693	1597	0		
Bifidobacterium animalis animalis YL2	Finished	University of Berne	2721755707	1800480	1520	0		
Bifidobacterium bifidum BGN4	Finished	Korea Research Institute of Bioscience and Biotechnology	2526164719	2223664	1896	0		
Bifidobacterium bifidum CBT BF3	Finished	DNA Link Inc.	2747842551	2210370	1847	0		
Bifidobacterium bifidum PRL2010	Finished	University of Parma	649633015	2214656	1767	0		
Bifidobacterium bifidum S17	Finished	Eurofins Medigenomix GmbH	649633016	2186882	1845	0		
Bifidobacterium breve 12L	Finished	University of Parma	2565956586	2244624	1821	0		
Bifidobacterium breve 689b	Finished	University of Parma	2565956585	2331707	1879	0		
Bifidobacterium breve ACS-071-V-Sch8b	Finished	J. Craig Venter Institute (JCVI)	651053005	2327492	1890	0		
Bifidobacterium breve CBT BR3	Finished	DNA Link Inc.	2747842549	2426006	2194	0		
Bifidobacterium breve JCM 7017	Finished	University College Cork	2565956584	2288919	1828	0		
Bifidobacterium breve JCM 7019	Finished	University College Cork	2588253744	2359009	1976	0		
Bifidobacterium breve LMG520	Finished	Korea University	2744055085	2403402	2127	0		
Bifidobacterium breve NCFB 2258	Finished	University College Cork	2565956583	2315904	1892	0		
Bifidobacterium breve S27	Finished	Eurofins MWG Operon	2561511061	2294458	1809	0		
Bifidobacterium breve UCC2003 (NCIMB8807)	Finished	University College Cork	651053006	2422684	1914	0		
Bifidobacterium catenulatum DSM 16992	Permanent Draft	Inner Mongolia Agricultural University	2660237904	2077949	1843	0		
Bifidobacterium catenulatum DSM 16992, JCM 1194, LMG 11043	Permanent Draft	Washington University in St. Louis	642979312	2058429	2009	0		
Bifidobacterium catenulatum DSM 16992, JCM 1194, LMG 11043	Permanent Draft	University of Parma	2597490243	2082756	1797	0		
Bifidobacterium choerinum DSM 20434	Permanent Draft	Inner Mongolia Agricultural University	2660238656	2024990	1696	0		
Bifidobacterium choerinum DSM 20434	Permanent Draft	DOE Joint Genome Institute (JGI)	2529292584	2037324	1671	0		
Bifidobacterium choerinum LMG 10510	Permanent Draft	University of Parma	2597490261	2096121	1752	0		
Bifidobacterium dentium ATCC 27678	Permanent Draft	Washington University in St. Louis	641736189	2642081	2499	0		
Bifidobacterium dentium Bd1	Finished	National University of Ireland, University of Parma, Agencourt Bioscience	646311910	2636367	2197	0		
Bifidobacterium dentium DSM 20436	Permanent Draft	DOE Joint Genome Institute (JGI)	2634166294	2668067	2248	0		
Bifidobacterium dentium JCVIHMP022	Permanent Draft	J. Craig Venter Institute (JCVI)	649989915	2636584	2341	0		
Bifidobacterium gallicum DSM 20093	Permanent Draft	Inner Mongolia Agricultural University	2671180475	1983163	1601	0		
Bifidobacterium gallicum DSM 20093, LMG 11596	Permanent Draft	Washington University in St. Louis	2562617144	2019802	1979	0		
Bifidobacterium gallicum DSM 20093, LMG 11596	Permanent Draft	University of Parma	2597490252	2004594	1603	0		
Bifidobacterium longum infantis 157F-NC	Finished	University of Tokyo	649633017	2408831	2070	0		
Bifidobacterium longum infantis CBT BT1	Finished	DNA Link Inc.	2747842548	2578115	2279	0		
Bifidobacterium longum infantis CECT 7210	Finished	Laboratorios Ordesa	2654587838	2467698	2162	0		
Bifidobacterium longum infantis JCM 1222	Finished	University of Tokyo	651053007	2828958	2641	0		
Bifidobacterium longum longum AH1206	Finished	University College Cork	2721755679	2421287	2144	0		
Bifidobacterium longum longum BBMN68	Finished	Beijing Genomics Institute (BGI)	649633018	2265943	1870	0		
Bifidobacterium longum longum CCUG30698	Finished	University College Cork	2687453622	2458004	2153	0		
Bifidobacterium longum longum DSM 20219	Permanent Draft	DOE Joint Genome Institute (JGI)	2634166334	2449019	2105	0		
Bifidobacterium longum longum F8	Finished	Wellcome Trust Sanger Institute	650377912	2384987	1744	0		
Bifidobacterium longum longum GT15	Finished	Vavilov Institute of General Genetics	2630968552	2337521	1980	0		
Bifidobacterium longum longum JCM 1217	Finished	University of Tokyo	649633019	2385164	2009	0		
Bifidobacterium longum longum JDM301	Finished	Shanghai Jiao Tong University	646564512	2477838	2022	0		
Bifidobacterium longum longum KACC 91563	Finished	National Academy of Agricultural Science, RDA, Korea	651053008	2395764	2050	0		
Bifidobacterium longum longum NCIMB8809	Finished	University College Cork	2687453409	2340989	1981	0		
Bifidobacterium longum sv. Infantis	Finished	DOE Joint Genome Institute (JGI)	643348516	2832748	2577	0		
Bifidobacterium pseudocatenulatum DSM 20438	Permanent Draft	Inner Mongolia Agricultural University	2695420736	2263778	1884	0		
Bifidobacterium pseudocatenulatum DSM 20438, JCM 1200, LMG 10505	Permanent Draft	Washington University in St. Louis	2562617145	2304808	2219	0		
Bifidobacterium pseudocatenulatum DSM 20438, JCM 1200, LMG 10505	Permanent Draft	University of Parma	2597490216	2283768	1882	0		
Bifidobacterium ruminantium DSM 6489	Permanent Draft	DOE Joint Genome Institute (JGI)	2558860117	2218565	1897	0		
Bifidobacterium ruminantium DSM 6489	Permanent Draft	Inner Mongolia Agricultural University	2657244954	2222186	1901	0		
Bifidobacterium ruminantium LMG 21811	Permanent Draft	University of Parma	2597490258	2249806	1922	0		

Marinelli *et al.* Table 2		
Species	**References**	**4.1.99.1 (KEGG)**
Bifidobacterium adolescentis	no indole (Zinedine&Faid 2007)	No
Bifidobacterium angulatum		No
Bifidobacterium animalis ssp. animalis		No
Bifidobacterium bifidum	no indole (Zinedine&Faid 2007)	No
Bifidobacterium breve		No
Bifidobacterium catenulatum		No
Bifidobacterium choerinum		
Bifidobacterium dentium		No
Bifidobacterium longum subsp. longum	no indole (Zinedine&Faid 2007)	No
Bifidobacterium longum subsp. infantis	no indole (Zinedine&Faid 2007)	
Bifidobacterium pseudocatenulatum	no indole (Zinedine&Faid 2007)	No

Marinelli *et al.* Suplementary Table 1

Espece	Aerobic	Phylum	Order	Genus	Phylogenetic group	OD_{600}	Medium	Acetate (mM)	AhR activity
Bifidobacterium adolescentis DSM20083, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. adolescentis	0,88	M58	73,420	1,32
Bifidobacterium adolescentis DSM20086	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. adolescentis	2,00	M58	NA	1,64
Bifidobacterium adolescentis DSM20086	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. adolescentis	1,48	LYBHI 4 JR	NA	1,95
Bifidobacterium angulatum DSM20098, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. adolescentis	1,07	M58	69,210	1,19
Bifidobacterium animalis subsp. animalis DSM20104, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. pseudolongum	0,68	M58	49,850	1,56
Bifidobacterium animalis subsp. animalis DSM20104, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. pseudolongum	NA	M58	NA	1,64
Bifidobacterium bifidum DSM20082	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	NA	0,79	M58	69,390	1,31
Bifidobacterium bifidum DSM20082	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	NA	NA	M58	NA	1,42
Bifidobacterium breve ATCC15701	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. breve and B. longum	0,31	M104	15,610	1,71
Bifidobacterium breve ATCC15701	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. breve and B. longum	NA	M58	NA	1,89
Bifidobacterium breve DSM20091	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. breve and B. longum	0,52	M104	1,130	1,04
Bifidobacterium catenulatum DSM16992, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. adolescentis	1,07	M58	62,140	1,33
Bifidobacterium choerinum DSM20434, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. pseudolongum	NA	M58	NA	1,68
Bifidobacterium choerinum DSM20434, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. pseudolongum	0,97	M58	69,050	1,84
Bifidobacterium dentium DSM20084	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. adolescentis	0,88	M104	48,720	1,18
Bifidobacterium dentium DSM20436, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. adolescentis	0,80	M104	42,870	1,78
Bifidobacterium gallicum DSM20093, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. pseudolongum	0,58	M58	51,140	1,09
Bifidobacterium longum subsp. infantis DSM20088, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. breve and B. longum	0,76	M58	49,770	1,37
Bifidobacterium longum subsp. infantis DSM20088, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. breve and B. longum	1,32	M58	NA	2,17
Bifidobacterium longum subsp. longum DSM20219, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. breve and B. longum	0,90	M58	65,040	1,48
Bifidobacterium longum subsp. longum DSM20219, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. breve and B. longum	NA	M58	NA	2,27
Bifidobacterium pseudocatenulatum DSM20438, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. adolescentis	1,27	M58	71,020	1,26
Bifidobacterium ruminantium DSM6489, Type strain	No	Actinobacteria	Bifidobacteriales	Bifidobacterium	B. adolescentis	1,08	M58	61,090	1,14

Marinelli *et al.* Suplementary Table 2

Short name	Name	Origin	Composition
BHI aerobic	Brain Heart Infusion	BD Difco	
LYBHI	Liquid Yeast Brain Infusion	Homemade	BHI 37g/L+ Yeast extact 5g/ml+ Hemin 0.1% 10mL/L+ Resazurin 0.1% 1ml/L
LYBHI 1	Liquid Yeast Brain Heart Infusion 1	Homemade	LYBHI+ Cystein HCL 0.6g/L
LYBHI 4	Liquid Yeast Brain Heart Infusion 4	Homemade	LYBHI+ Cystein HCL 0.6g/L+ cellobiose 0.5g/L+ maltose 0.5g/L+ soluble starch 0.5g/L
LYBHI 4 RJ	Liquid Yeast Brain Heart Infusion 4 Rumen Juice	Homemade	LYBHI+ Cystein HCL 0.6g/L+ cellobiose 0.5g/L+ maltose 0.5g/L+ starch 0.5g/L+ rumen fluid 10% (v/v)
M104	DSMZ PYG MEDIUM + K1 Vitamin	Homemade	
M58	DSMZ BIFIDOBACTERIUM MEDIUM	Homemade	

Alexeev, E. E., Lanis, J. M., Kao, D. J., Campbell, E. L., Kelly, C. J., Battista, K. D., … Colgan, S. P. (2018). Microbiota-Derived Indole Metabolites Promote Human and Murine Intestinal Homeostasis through Regulation of Interleukin-10 Receptor. *The American Journal of Pathology*. https://doi.org/10.1016/j.ajpath.2018.01.011

Bansal, T., Alaniz, R. C., Wood, T. K., & Jayaraman, A. (2010). The bacterial signal indole increases epithelial-cell tight-junction resistance and attenuates indicators of inflammation. *Proceedings of the National Academy of Sciences*, *107*(1), 228–233. https://doi.org/10.1073/pnas.0906112107

Biagi, E., Candela, M., Turroni, S., Garagnani, P., Franceschi, C., & Brigidi, P. (2013). Ageing and gut microbes: Perspectives for health maintenance and longevity. *Pharmacological Research*, *69*(1), 11–20. https://doi.org/10.1016/j.phrs.2012.10.005

Biavati, B., & Mattarelli, P. (2006). The Family Bifidobacteriaceae. In *The Prokaryotes* (pp. 322–382). New York, NY: Springer New York. https://doi.org/10.1007/0-387-30743-5_17

Cerf-Bensussan, N., & Gaboriau-Routhiau, V. (2010, October 1). The immune system and the gut microbiota: Friends or foes? *Nature Reviews Immunology*. Nature Publishing Group. https://doi.org/10.1038/nri2850

De Vuyst, L., & Leroy, F. (2011). Cross-feeding between bifidobacteria and butyrate-producing colon bacteria explains bifdobacterial competitiveness, butyrate production, and gas production. *International Journal of Food Microbiology*, *149*(1), 73–80. https://doi.org/10.1016/j.ijfoodmicro.2011.03.003

Di Gioia, D., Aloisio, I., Mazzola, G., & Biavati, B. (2014). Bifidobacteria: their impact on gut microbiota composition and their applications as probiotics in infants. *Applied Microbiology and Biotechnology*, *98*(2), 563–577. https://doi.org/10.1007/s00253-013-5405-9

Ewaschuk, J. B., Diaz, H., Meddings, L., Diederichs, B., Dmytrash, A., Backer, J., … Madsen, K. L. (2008). Secreted bioactive factors from *Bifidobacterium infantis* enhance epithelial cell barrier function. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, *295*(5), G1025–G1034. https://doi.org/10.1152/ajpgi.90227.2008

Flaveny, C. A., Murray, I. A., Chiaro, C. R., & Perdew, G. H. (2009). Ligand Selectivity and Gene Regulation by the Human Aryl Hydrocarbon Receptor in Transgenic Mice. *Molecular Pharmacology*, *75*(6), 1412–1420. https://doi.org/10.1124/mol.109.054825

Fujii-Kuriyama, Y., Ema, M., Mimura, J., & Sogawa, K. (1994). Ah receptor: a novel ligand-activated transcription factor. *Experimental and Clinical Immunogenetics*, *11*(2–3), 65–74. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7826667

Fukuda, S., Toh, H., Hase, K., Oshima, K., Nakanishi, Y., Yoshimura, K., … Ohno, H. (2011). Bifidobacteria can protect from enteropathogenic infection through production of acetate. *Nature*, *469*(7331), 543–549. https://doi.org/10.1038/nature09646

Fukumoto, S., Toshimitsu, T., Matsuoka, S., Maruyama, A., Oh-oka, K., Takamura, T., … Nakao, A. (2014). Identification of a probiotic bacteria-derived activator of the aryl hydrocarbon receptor that inhibits colitis. *Immunology and Cell Biology*, *92*(5), 460–5. https://doi.org/10.1038/icb.2014.2

Gagliani, N., Vesely, M. C. A., Iseppon, A., Brockmann, L., Xu, H., Palm, N. W., … Flavell, R. a. (2015). Th17 cells transdifferentiate into regulatory T cells during resolution of inflammation. *Nature*, 1–5. https://doi.org/10.1038/nature14452

Grimm, V., Westermann, C., & Riedel, C. U. (2014). Bifidobacteria-host interactions--an update on colonisation factors. *BioMed Research International*, *2014*, 960826. https://doi.org/10.1155/2014/960826

Hankinson, O. (1995). The aryl hydrocarbon receptor complex. *Annual Review of Pharmacology and Toxicology*, *35*(1), 307–340. https://doi.org/10.1146/annurev.pa.35.040195.001515

Hashimoto, T., Perlot, T., Rehman, A., Trichereau, J., Ishiguro, H., Paolino, M., … Penninger, J. M. (2012). ACE2 links amino acid malnutrition to microbial ecology and intestinal inflammation. *Nature*, *487*, 477–483. https://doi.org/10.1038/nature11228

Hsieh, C.-Y., Osaka, T., Moriyama, E., Date, Y., Kikuchi, J., & Tsuneda, S. (2015). Strengthening of the intestinal epithelial tight junction by *Bifidobacterium bifidum*. *Physiological Reports*, *3*(3), e12327. https://doi.org/10.14814/phy2.12327

Hubbard, T. D., Murray, I. A., Bisson, W. H., Lahoti, T. S., Gowda, K., Amin, S. G., … Perdew, G. H. (2015). Adaptation of the human aryl hydrocarbon receptor to sense microbiota-derived indoles.

Scientific Reports, *5*(1), 12689. https://doi.org/10.1038/srep12689

Hubbard, T. D., Murray, I. A., & Perdew, G. H. (2015, September 1). Special section on drug metabolism and the microbiome - Minireview indole and tryptophan metabolism: Endogenous and dietary routes to ah receptor activation. *Drug Metabolism and Disposition*. https://doi.org/10.1124/dmd.115.064246

Iwabuchi, N., Takahashi, N., Xiao, J., Miyaji, K., & Iwatsuki, K. (2007). In vitro Th1 cytokine-independent Th2 suppressive effects of bifidobacteria. *Microbiology and Immunology*, *51*(7), 649–60. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17641467

Jeon, S. G., Kayama, H., Ueda, Y., Takahashi, T., Asahara, T., Tsuji, H., … Takeda, K. (2012). Probiotic Bifidobacterium breve Induces IL-10-Producing Tr1 Cells in the Colon. *PLoS Pathogens*, *8*(5), e1002714. https://doi.org/10.1371/journal.ppat.1002714

Jin, U.-H., Lee, S.-O., Sridharan, G., Lee, K., Davidson, L. A., Jayaraman, A., … Safe, S. (2014). Microbiome-Derived Tryptophan Metabolites and Their Aryl Hydrocarbon Receptor-Dependent Agonist and Antagonist Activities. *Molecular Pharmacology*, *85*(May), 777–788. https://doi.org/10.1124/mol.113.091165

Lamas, B., Richard, M. L., Leducq, V., Pham, H.-P., Michel, M.-L., Da Costa, G., … Sokol, H. (2016). CARD9 impacts colitis by altering gut microbiota metabolism of tryptophan into aryl hydrocarbon receptor ligands. *Nature Medicine*, *22*(6), 598–605. https://doi.org/10.1038/nm.4102

Lee, J. S., Cella, M., McDonald, K. G., Garlanda, C., Kennedy, G. D., Nukaya, M., … Colonna, M. (2012). AHR drives the development of gut ILC22 cells and postnatal lymphoid tissues via pathways dependent on and independent of Notch. *Nat Immunol*, *13*(2), 144–151. https://doi.org/10.1038/ni.2187

Liu, Z., Li, L., Chen, W., Wang, Q., Xiao, W., Ma, Y., … Yang, H. (2018). Aryl hydrocarbon receptor activation maintained the intestinal epithelial barrier function through Notch1 dependent signaling pathway. *International Journal of Molecular Medicine*, *41*(3), 1560–1572. https://doi.org/10.3892/ijmm.2017.3341

López, P., González-Rodríguez, I., Sánchez, B., Ruas-Madiedo, P., Suárez, A., Margolles, A., & Gueimonde, M. (2012). Interaction of Bifidobacterium bifidum LMG13195 with HT29 cells influences regulatory-T-cell-associated chemokine receptor expression. *Applied and Environmental Microbiology*, *78*(8), 2850–7. https://doi.org/10.1128/AEM.07581-11

Martinez, F. A. C., Balciunas, E. M., Converti, A., Cotter, P. D., & de Souza Oliveira, R. P. (2013). Bacteriocin production by Bifidobacterium spp. A review. *Biotechnology Advances*, *31*(4), 482–488. https://doi.org/10.1016/j.biotechadv.2013.01.010

Ménard, O., Butel, M.-J., Gaboriau-Routhiau, V., & Waligora-Dupriet, A.-J. (2008). Gnotobiotic mouse immune response induced by Bifidobacterium sp. strains isolated from infants. *Applied and Environmental Microbiology*, *74*(3), 660–6. https://doi.org/10.1128/AEM.01261-07

Monteleone, I., Pallone, F., & Monteleone, G. (2013). Aryl hydrocarbon receptor and colitis. *Seminars in Immunopathology*, *35*(6), 671–675. https://doi.org/10.1007/s00281-013-0396-2

Monteleone, I., Rizzo, A., Sarra, M., Sica, G., Sileri, P., Biancone, L., … Monteleone, G. (2011). Aryl Hydrocarbon Receptor-Induced Signals Up-regulate IL-22 Production and Inhibit Inflammation in the Gastrointestinal Tract. *Gastroenterology*, *141*(1), 237–248.e1. https://doi.org/10.1053/j.gastro.2011.04.007

Murray, I. A., & Perdew, G. H. (2017). Ligand activation of the Ah receptor contributes to gastrointestinal homeostasis. *Current Opinion in Toxicology*, *2*, 15–23. https://doi.org/10.1016/j.cotox.2017.01.003

Quintana, F. J., & Sherr, D. H. (2013). Aryl hydrocarbon receptor control of adaptive immunity. *Pharmacological Reviews*. https://doi.org/10.1124/pr.113.007823

Raimondi, S., Amaretti, A., Leonardi, A., Quartieri, A., Gozzoli, C., & Rossi, M. (2016). Conjugated Linoleic Acid Production by Bifidobacteria: Screening, Kinetic, and Composition. *BioMed Research International*, *2016*, 8654317. https://doi.org/10.1155/2016/8654317

Rivière, A., Selak, M., Lantin, D., Leroy, F., & De Vuyst, L. (2016). Bifidobacteria and Butyrate-Producing Colon Bacteria: Importance and Strategies for Their Stimulation in the Human Gut. *Frontiers in Microbiology*, *7*, 979. https://doi.org/10.3389/fmicb.2016.00979

S. Ranhotra, H. (2017). Gut Microbiota and Host Nuclear Receptors Signalling. *Nuclear Receptor*

Research, *4*. https://doi.org/10.11131/2017/101316

Schiering, C., Vonk, A., Das, S., Stockinger, B., & Wincent, E. (2018). Cytochrome P4501-inhibiting chemicals amplify aryl hydrocarbon receptor activation and IL-22 production in T helper 17 cells. *Biochemical Pharmacology*, *151*, 47–58. https://doi.org/10.1016/j.bcp.2018.02.031

Shimada, Y., Kinoshita, M., Harada, K., Mizutani, M., Masahata, K., Kayama, H., & Takeda, K. (2013). Commensal bacteria-dependent indole production enhances epithelial barrier function in the colon. *PLoS ONE*, *8*(11), e80604. https://doi.org/10.1371/journal.pone.0080604

Spiljar, M., Merkler, D., & Trajkovski, M. (2017). The Immune System Bridges the Gut Microbiota with Systemic Energy Homeostasis: Focus on TLRs, Mucosal Barrier, and SCFAs. *Frontiers in Immunology*, *8*, 1353. https://doi.org/10.3389/fimmu.2017.01353

Stockinger, B., Meglio, P. Di, Gialitakis, M., & Duarte, J. H. (2014). The Aryl Hydrocarbon Receptor: Multitasking in the Immune System. *Annual Review of Immunology*, *32*(1), 403–432. https://doi.org/10.1146/annurev-immunol-032713-120245

Voreades, N., Kozil, A., & Weir, T. L. (2014). Diet and the development of the human intestinal microbiome. *Frontiers in Microbiology*, *5*, 494. https://doi.org/10.3389/fmicb.2014.00494

Wells, J. M., Rossi, O., Meijerink, M., & van Baarlen, P. (2011). Epithelial crosstalk at the microbiota-mucosal interface. *Proceedings of the National Academy of Sciences of the United States of America*, *108 Suppl*(Supplement_1), 4607–4614. https://doi.org/10.1073/pnas.1000092107

Yu, M., Wang, Q., Ma, Y., Li, L., Yu, K., Zhang, Z., … Yang, H. (2018). Aryl Hydrocarbon Receptor Activation Modulates Intestinal Epithelial Barrier Function by Maintaining Tight Junction Integrity. *International Journal of Biological Sciences*, *14*(1), 69–77. https://doi.org/10.7150/ijbs.22259

Yu, S., & Gao, N. (2015). Compartmentalizing intestinal epithelial cell toll-like receptors for immune surveillance. *Cellular and Molecular Life Sciences*, *72*(17), 3343–3353. https://doi.org/10.1007/s00018-015-1931-1

Zelante, T., Iannitti, R. G., Cunha, C., DeLuca, A., Giovannini, G., Pieraccini, G., … Manuscript, A. (2013). Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22. *Immunity*, *39*(6), 372–385. https://doi.org/10.1016/j.immuni.2013.08.003

Zelante, T., Pieraccini, G., Scaringi, L., Aversa, F., & Romani, L. (2016, March). Learning from other diseases: protection and pathology in chronic fungal infections. *Seminars in Immunopathology*. https://doi.org/10.1007/s00281-015-0523-3

Chapter 4. Discussion

The gastrointestinal tract is a large absorptive surface continuously challenged by environmental stimuli, such as food, xenobiotic, pathogens and metabolites originated from the commensal microbiota. In the last decades, the knowledge on the complex interaction between the human and the residing microbial community evolved, encouraging the scientific efforts toward the description of cellular mechanisms involved in this delicate dialogue and the physiological effects. Several cell receptors have been identified and described for their role in mediating the bacterial signals into an effect on the host wellbeing. Among these receptors, it became evident that the aryl hydrocarbon receptor (AhR) is important in sensing the environmental signal, among which bacterial metabolites. Able to bind molecules of different origins, AhR has been described to sense metabolites derived from pathogens as well as from commensal bacteria and stimulate a general local and systemic protective effect. The main described microbial-derived AhR ligands are indoles, produced by some commensal and probiotic bacteria. By ligand-binding to AhR, indoles have been described to increase tight junction integrity, reduced inflammation and control of pathogen infections in the gut (Bansal et al. 2010; Zelante et al. 2013; Lamas et al. 2018). However, despite the increasing interest in the microbiota-AhR interaction, the knowledge on bacterial-produced AhR ligands, other than indoles and indole-derivatives, is still limited. Considering the abundance and the extensive metabolic capacity of the gut microbiota, it is likely that other bacterial metabolites, are produced and present in the intestine, at concentrations able to stimulate the AhR signalling pathway that might impact the host physiology.

In this context, the two presented papers aimed to identify novel microbial species and metabolites activating the AhR signalling pathway in human intestinal epithelial cells. By screening a collection of 132 gut commensal bacterial strains (106 different species) on an AhR reporter system expressed in human intestinal epithelial cell lines, we identified some bacteria activating the AhR signalling pathway. The quantification of main organic acids produced in the bacterial cultures, allowed the description of two main groups of activator bacteria. A first group of butyrate-producers made the object of the first paper, while the second paper was focused on a group of activator bacteria not producers of short-chain fatty acids (SCFAs) nor indoles.

In the first paper we confirmed that part of tested bacteria activates AhR signalling through the production of butyrate and that this effect is reproducible in another intestinal epithelial cell line (Caco-2). Incubating intestinal epithelial cell lines in presence of butyrate we demonstrated that this SCFA stimulates the expression of the AhR dependent genes *CYP1A1*, *AhR* and *AhRR* as well as the AHR protein expression. The induced AhR activity, evidenced through the reporter system, and the induction of genes and protein expressions, suggested a direct role of the butyrate on the AhR signalling pathway.

Considering that the regulation of gene transcription by butyrate is described to involve a wide range of transcription factors (TF), we performed a preliminary analysis of human AhR promoter. As expected from previous publications, we found several binding sites for TF implicated in butyrate-regulated gene expression including Specificity Protein-1 (SP1) binding GC-rich boxes, Activator protein 1 (AP-1) and NFκB responsive elements. In order to investigate if these TF are targeted by butyrate and whether they directly affect *AhR* expression and the consequent activation of the AhR reporter system, we inhibited these TF in HT29-AhR cell line in presence and absence of butyrate.

By inhibiting Sp-1 we observed a down-regulation of the AhR induced response by the prototypical ligand TCDD, which is described to activate AhR by an exclusively ligand-binding mechanism (*Figure 1*). This suggest that Sp1 has an impact on the basal AHR protein expression and, consequently, on the AhR activity. Therefore, the inhibition of butyrate-induced AhR response could not be assumed as relevant. Our observation is coherent with previous reports describing that the basal AhR expression is regulated by the expression and distribution of Sp1-like factors (Fitzgerald et al. 1998), so we could not exclude nor confirm the involvement of Sp1 in the butyrate-induced activation of AhR in our experimental conditions.

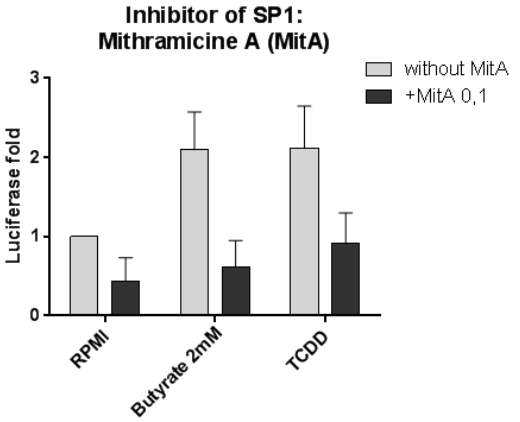


Figure 348: *Butyrate incubated in presence and absence of the Sp-1 inhibitor (Mithramycine A - MitA) in HT29-AhR cell lines.*

We then tested the effects of a NFκB inhibitor (BAY-117082) and a proteasome inhibitor (MG-132) (Figure 2A) as well as an AP-1 inhibitor (SR-11302) (Figure 2B) in presence or absence of butyrate. Our results showed that the inhibition of both NFκB and AP-1 does not affect the

butyrate-induced AhR activation, suggesting that these TF are not targeted by butyrate for inducing the described AhR activation.

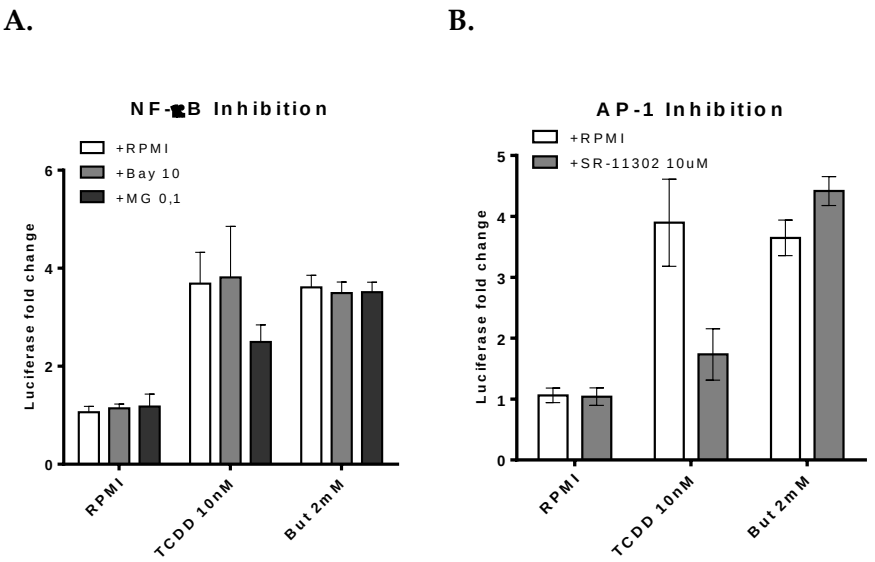


Figure 359: *Butyrate incubated in presence and absence of a NFκB inhibitor (BAY-117082) and a proteasome inhibitor (MG-132)(A) and an AP-1 inhibitor (SR-11302)(B) in HT29-AhR cell line*

Overall these preliminary results (not shown in the submitted manuscript) demonstrated that the butyrate-induced activation of AhR is not mediated by TF present in the promoter region of AhR, in our conditions, thus opening the way to the implication of other possible targets for butyrate directly involved in the mechanism.

Butyrate, like other SCFAs, activates eukaryotic cells through two main mechanisms: activation of specific G-protein coupled receptors (GPCR: GPR41, GPR43, GPR109a) and inhibition of histone deacetylases (HDAC). Considering that the three G-protein coupled receptors are expressed in HT-29 and Caco-2 cell lines (Martin-Gallausiaux et al. n.d.), we tested the hypothesis that the AhR butyrate-induced activation could be mediated by GPRs signalling. However, as other tested GPR agonists did not induced activates AhR signalling in our reporter system, we excluded that these coupled receptor are involved in the direct effect induced by butyrate.

SCFAs, *via* their ability to inhibit HDAC, are potent modulators of histones and TF acetylation, which are well-documented regulatory mechanisms of gene transcription. A recent study showed that SCFAs and other HDAC inhibitors (HDACi), enhanced AhR-induced genes such

as *CYP1A1 via* the increase of histone acetylation (Jin et al. 2017). These results prompted us to investigate the role of HDACi in AhR activation. At transcriptional level, we confirmed the role of HDACi, included butyrate, in inducing the expression of *CYP1A1*, *AHRR* and *AhR* genes. However, we also highlight a second role of the butyrate at cellular level, independent on its role as HDACi. Indeed, the activation observed on the reporter system only with butyrate and not with other HDACi, suggested a role on the protein translocation and expression, that we confirmed by immunoblotting assay. Thus, we identified two distinct effects at transcriptional level to explain the butyrate-induced activation of the AhR-dependent genes: a first role at epigenetic level on the histone acetylation and a second role involving the activation of the signalling pathway through the nuclear translocation of the receptor. Therefore, to explain the induced activation of AhR signalling by butyrate, independently of its role as HDACi, we wondered if it could exert a role as ligand.

Up to date, the main AhR described ligands are characterized by the presence of aromatic rings. In my opinion, this peculiarity could be explained by the fact that many of the studies that aimed to identify new AhR ligands, predicted the potential active molecule starting from the described AhR agonists. Therefore, considering the strong interest of toxicologists in the AhR signalling since its discovery, it seems not surprisingly that most of the predicted ligand shared the aromatic ring characteristic of the first described classes of AhR ligands (HAHs and PAHs). However, in order to sense environmental signals, including microbiota and pathogens, I suppose that the ability of AhR to bind endogenous and exogenous molecules is likely to extend toward diverse molecular structures.

A novel mechanism for the butyrate to impact the host has been described, through the direct ligand binding to the nuclear receptor proliferator-activated receptor-γ (PPARγ) (Sheril Alex et al. 2013). Similarly to AhR, PPARγ is a well-characterized receptor targeting endogenous and exogenous ligands. This newly described role of butyrate in ligand binding to PPARγ, prompts us to speculate about a possible similar role on AhR. Indeed, the inhibition of ligand binding through specific AhR antagonists, resulted in a significant decrease in the butyrate-induced activation in reporter system, suggesting a mechanism dependent on ligand binding to the cytoplasmatic receptor. However, the milder inhibitory effect evidenced at gene level in presence of the AhR antagonists, could be explained by the epigenetic impact of butyrate as HDACi on the expression of AhR-dependent genes.

The ligand binding to AhR is described to induce the translocation of the receptor in the nucleus, to ensure the formation of the AhR:ARNT complex and the consequent interaction with XRE motifs. Coherently to the hypothesis of a ligand-dependent activation, we described a redistribution of AHR protein from the cytoplasm into the nucleus upon treatment with butyrate. Upon nuclear translocation, the HSP90 displacement occurs in order to allow the formation of AHR:ARNT complex and the consequent binding to XRE motifs. By inhibiting the displacement of HSP90 (with EGCG) upon nuclear shuttling, we observed a reduction in butyrate-induced AhR activation, coherent with what observed for TCDD. This indirect evidence could suggest that the butyrate-activation of AhR signalling is dependent on the AHR:ARNT complex formation and the XRE interaction, to stimulate a transcriptional effect. However, other investigation are necessary to explain the implication of the butyrate on the AhR:ARNT:XRE complex. Lastly, to find additional supports for butyrate behaving as selective AhR ligand, we performed structural modelling of the binding of butyrate to human AhR (hAhR) demonstrating a likely binding of this SCFA in the binding pocket, in four possible conformations. To find additional evidences to our hypothesis, a direct assessment of the binding by competitive binding assay could have provided additional supports to our results, in particular concerning the binding affinity of the butyrate for AhR.

These results suggested that butyrate acts as an AhR ligand which is, to our knowledge, an original mechanism only been reported for another ligand-binding transcription factor, PPARγ (Sheril Alex et al. 2013). Based on this conclusion, we proposed a mechanism through which butyrate, absorbed by IECs in intestinal lumen, binds to the cytoplasmic AhR complex and stimulates a conformational change leading to the translocation of the receptor in the nucleus. Then, through the formation of AhR:ARNT:XRE, the butyrate stimulates the expression of AhR-dependent genes, among which *CYP1A1* and *AHRR* and drive the expression of AHR protein. Additionally, the expression of AhR-dependent genes induced by the activation signalling seem to be further supported by the role of butyrate as HDACi.

The description of the mechanism might have been further detailed by studying the role of butyrate in the formation of the AHR:ARNT:XRE complex. Thus, electrophoretic mobility shift assay could have shed light on the impact of butyrate on the AhR-XRE binding and better decipher the activation pathway.

The identification of butyrate as an AhR ligand open the way to different researches field. On one side the description that AhR binds to a so small molecule and so structurally different from the previously described agonists, might lead to the identification of new classes of AhR ligands and potential new roles for other SCFA. In this regard, other SCFA emerged to activate the AhR signalling, such as propionate, valerate, isovalerate and isobutyrate. Despite the fact that we did not investigate their mechanism for the AhR activation, these SCFA could be considered as potentially targeting AhR. To my knowledge, a little is known about the impact of valerate, isovalerate and isobutyrate on other transcription factor and it remains unclear whether their concentration in the gut could be enough to stimulate AhR signalling at physiologically conditions. Conversely, propionate has been largely described for its physiological roles in the gut. Similarly to butyrate, propionate activates the transcription activity of PPARα (Higashimura et al. 2015) and PPARγ (Nepelska et al. 2017; Sheril Alex et al. 2013). By demonstrating the ligand-mediated PPARγ activation by butyrate, Alex and co-workers (Sheril Alex et al. 2013) showed similar transcriptional effects induced by propionate that could suggest an analogous role of propionate, although not proposed by the authors. Coherently, considering the similarity in inducing PPARγ and AhR we suspect that propionate could be another promising modulator of AhR signalling, but supplementary studies are necessary to decipher the cellular mechanism.

A second question that could emerged from the evidenced ability of butyrate to bind to both AhR and PPARγ is whether it binds to other ligand-activated transcription factors. To my knowledge no other studies have been conducted on the binding of butyrate on other ligand-activated transcription factor, but it could be an interesting point to further extend the roles of butyrate at cellular level.

A third question that emerged from our results is whether the butyrate stimulates a prolonged or transient stimulation of AhR. In the gut, the intensive production of butyrate by the microbiota and the continuous intake by enterocytes, could raise the issue on whether this SCFA could induced a transient or prolonged AhR stimulation. It is well known that up to 60-70% of the colonic energy need arise from butyrate (Schilderink 2013), so we could speculate that the eukaryotic cells will favour this essential metabolism, limiting the butyrate availability for other pathway, including AhR. Considering the structurally difference of butyrate compared to other AhR ligands, it is hard to imagine which could be its fate in the AhR

signalling, following activation. However, I would rather propose an induced transient stimulation that, to our knowledge, induce protective effects coherent with the described responses induced by butyrate on intestinal homeostasis, beyond the AhR signalling. Another issue about the microbial production of SCFA concerns the combined effect of different SCFA on the receptor. At present we do not have any information about a concerted effect of more SCFA on AhR signalling or in any others signalling but considering the dimension of the binding pocket and the possibility for arranging four butyrate molecules (as evidenced by our docking), acetate and propionate could likely find a favourable conformation along with butyrate. However, the effect of this concerted role remains an open question.

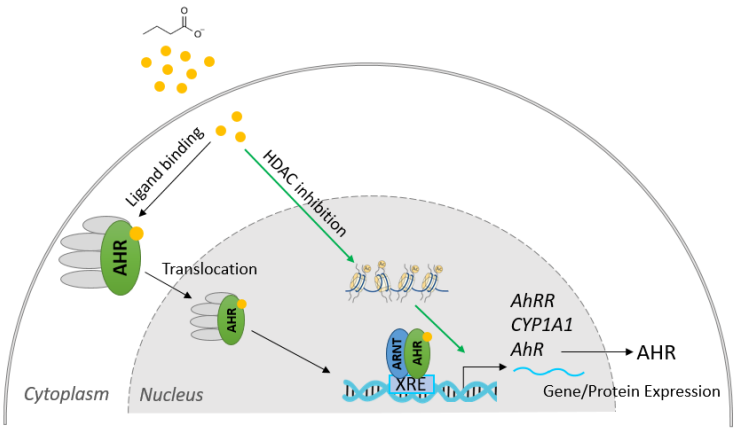


Figure 40: Schematic representation of the proposed mechanism for the AhR activation by butyrate

Beside the aforementioned SCFA-producing bacteria activating the AhR pathway, in the second paper we aimed to identify bacterial strains able to activate AhR independently from the production of SCFA. For this purpose, from the activator bacteria identified in the screening, we excluded the entire Phyla of Firmicutes and Fusobacteria, producers of butyrate, Bacteroidetes and the sole tested member of Verrucomicrobia (*Akkermansia muciniphila*), producers of propionate. Consequently the selection of SCFA-non-producing-bacteria and activating AhR, resulted in strains belonging to Proteobacteria and Actinobacteria, with the exception of *Propionibacteriales*, described to produce propionate. Consequently, we focused our attention on the Proteobacteria Phylum and on the Genus of *Bifidobacteriales* and *Eggerthellales*, belonging to Actinobacteria. Quantifying the organic acids produced by each bacterial culture, we confirmed the absence of butyrate and propionate production by Bifidobacteria. Additionally, the selected bacteria also showed no production of valerate,

isovalerate and isobutyrate, potentially activators of AhR signalling, as evidenced in our first paper. In order to identify bacteria activating AhR through the production of a novel active metabolite, we wanted to focus our attention on those not producers of indole, which has been already described as AhR ligand. Therefore, by cross-checking the bibliographic references with a preliminary comparative genomic analysis, we further reduced the list of selected promising bacteria by excluding those described or predicted for the production of indoles. Among the indole-producers, Proteobacteria have been described, so the evidenced AhR activation could be likely due to the production of this class of metabolites. We also excluded the sole member of *Eggerthellales* tested (*Adlercreutia equolifaciens*) because of the limited genomic and metabolic information available and annotated. The selection of tested bacteria, based on the AhR activation and metabolic activity (no production of SCFA and indole) results in the selection of *Bifidobacteriales.* Indeed, the better taxonomical and metabolic characterization of *Bifidobacteriales*, evidenced the absence of indole production, also confirmed by our preliminary analysis of the annotated genomes. Therefore *Bifidobacteriales* emerged as a promising group for the identification of a novel metabolite involved in a SCFA- and indole-independent activation of the AhR signalling pathway, which made the object of the second presented paper. Bifidobacteria is a well-studied genus of early colonizers bacteria that dominates the intestine of healthy breast-fed infant. In adults the level of Bifidobacteria is lower, compared to child, but stable. Numerous health-promoting effects have been ascribed to strains of the Bifidobacterium genus, among which the capacity to stimulate the immune system, particularly the T-helper 1(T_H1)/T_H2 balance, and the enhancement of intestinal epithelial barrier function, in particular tight junctions (Iwabuchi et al. 2007; Hsieh et al. 2015). These roles are generally mediated by the production of different metabolites, among which acetate, conjugate linoleic acid and bacteriocins. Thus, it is not surprising that Bifidobacteria are so widely used as probiotics in the treatment and prevention of intestinal diseases. In this context we wondered whether some of the health-promoting effects of Bifidobacteria could converge to some of the positive roles of AhR activation in intestinal epithelium.

In the second paper we demonstrated that some Bifidobacteria were able to activate the AhR reporter system in intestinal epithelial cell line. Not all of the tested bacteria showed the same AhR activation potential, which is coherent with reports describing a species-specific effect of Bifidobacteria on human intestinal homeostasis (Ménard et al. 2008). Additionally, we confirmed that some Bifidobacteria, selected among the activators, were able to induce the

expression of the AhR-dependent gene *CYP1A1*, confirming the role on the AhR signalling. Coherently, *B. breve* was shown to stimulate AhR gene expression in Tr1 cells (Jeon et al. 2012), suggesting that the Bifidobacteria-induced activation of the AhR signalling could involve cell populations other than IECs. To get insight the mechanism involved in the induced AhR activation, we hypothesized that some Bifidobacteria could produce metabolites acting as AhR ligands. Thus, by inhibiting the ligand binding through two specific AhR antagonists (CH-223191 and GNF-351), we evidenced a reduction in the AhR activity. The rich composition of the culture media has led a probable additional challenge to the discrimination among active molecules present in the non-inoculated media and bacterial produced-metabolites. However, the strong bacterial-induced activation and the consequent reduction in presence of ligand-binding antagonists, could not be completely explained, in our opinion, by the sole basal inhibition of the culture media. Consequently, we suspected that other possible ligands, different from those present in the media, may be responsible for the bacterial-induced activation of AhR signalling.

The described translocation of AhR in the nucleus induced by some of the selected Bifidobacteria, is coherent with the mechanism of other described AhR ligands. This provided an additional evidence supporting the hypothesis of a bacterial-produced ligand stimulating the AhR signalling pathway. Among metabolites produced by Bifidobacteria, acetate and lactate have been extensively studied for their impact on the host physiology. However, we did not evidenced any activation of the AhR reporter system induced by both acetate and lactate. Therefore, additional investigation is needed to identify the microbial-produced metabolites responsible for the AhR activation by Bifidobacteria.

Through the identification of Bifidobacteria as potential AhR activators in IECs, this second paper provided an additional support for the role of AhR as mediator of the human-microbiota crosstalk in intestine. Although, the identification of the involved metabolites as well as the complete involved mechanism remain only partly described in this work. The tools used to decipher the butyrate-dependent mechanism, could be used as well for describing the mechanism of the Bifidobacteria-induced activation. Starting from confirming that the induced activation is dependent on ligand-binding and on the nuclear translocation of the receptor, we could have had a way stronger argument for describing a Bifidobacteria-derived AhR ligands. Additionally, as for the butyrate, the deeper description of the activation

mechanism could have taken advantage from competitive binding and DNA binding assays. Lastly, a metabolomics analysis of the bacterial cultures could have provided fundamental support for the identification of the involved molecule. Coherently, the identification of a novel metabolite from some Bifidobacteria could open some question on the metabolism, genetics and taxonomy of the bacteria genus: the metabolism from which the active molecules is originated, the genes involved in this production and their expression among the *Bifidobacterium* genus as well as the impact for the host.

Considering the lack of metabolomics information about the nature of the active compound, we could speculate that its ligand-dependent activation of AhR may be originated in different ways: the direct bacterial production, the conversion of a precursor or a weak ligand in the medium into a strong ligand by Bifidobacteria or the concerted role of different molecules in the binding pocket. The first and most obvious hypothesis is the direct production of a novel AhR ligand by Bifidobacteria, but for its description, metabolomics and biochemical analysis are fundamental, along with a docking model of the identified molecule in the AhR binding pocket. However, coherently with previous evidences describing that weak ligands or ligand precursors could be converted in ligands with higher affinity, another hypothesis is the formation of the active molecule from precursors or weak ligands already present in the non-inoculated media. Indeed the basal activation observed from the non-inoculated media let suspect the presence of active metabolites prior to the bacterial inoculation. In this regard, the indole acetic acid, an AhR ligand, is the precursor of the microbial-produced skatole which is able to bind AhR and induce the signalling pathway, as well (Rasmussen et al. 2016). Another hypothesis for the observed activation is whether the acetate or lactate, produced by Bifidobacteria, could exert a concerted role in activating the receptor together with endogenous ligands. To our knowledge, the binding of acetate or lactate to other cellular receptors has not yet been addresses, thus the question on the possible effect remain open for future studies.

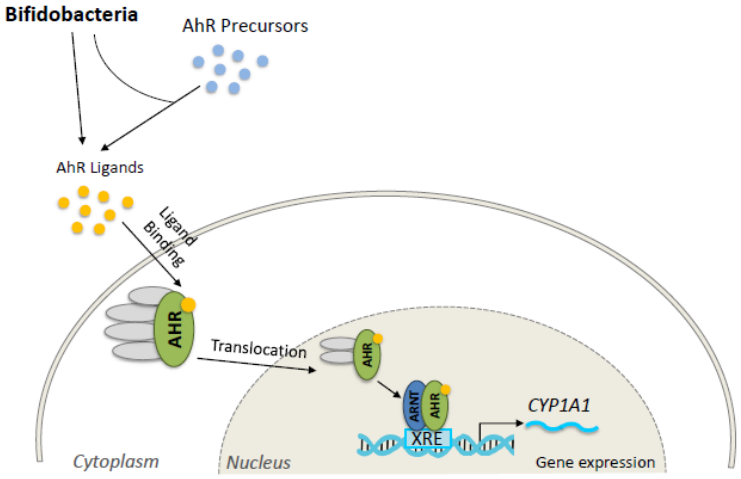


Figure 40: Schematic representation of the proposed mechanism for the AhR activation by Bifidobacteria.

Our cellular approach allowed the identification of one novel mechanism and a novel family of bacterial actors in the human-microbiota crosstalk through AhR. However, we did not investigate the possible implications of the AhR activation through these two pathways in the intestinal homeostasis. Thus, the role of butyrate as well the Bifidobacteria-derived metabolite via AhR on the intestinal epithelial layer, could be assessed on differentiated epithelial cells looking, for examples, at the tight junction formation and maintenance. However, the cellular approach is generally criticized when the obtained results are transposed to human physiology, because of the limitations of the cellular model. At the same time, considering the differences between human and murine AhR, we could not know, at the current state of art, if the AhR activation induced by butyrate as well as Bifidobacteria are human-specific or rather a mechanism shared between human and mouse. Preliminary tests have been performed to indirectly assess the role of butyrate in *Cyp1a1* gene expression in murine intestinal epithelial cell lines (ModeK and CMT-93). In our conditions and in both cell lines, we did not observed a butyrate-induced *Cyp1a1* gene expression and, even though other complementary assays are necessary to confirm this preliminary results, we supposed that the described ligand role of butyrate on AhR could be human-specific. The AhR reporter system expressed in both murine cell line could have been a useful tool for describing the role of butyrate on murine epithelial cells and ultimately identifying if the mechanism is human-specific or rather common to mice.

If we assume that the described mechanisms are specific to humans and not shared with mice, it became evident the limitation in using a murine model to highlight the physiological mechanism. However, the recently developed humanized mice, whose express human AhR, could be a valuable tool for studying possible human-specific AhR activations. Through this approach, *in vivo* effect of butyrate or microbial-derived ligand via AhR, could be more easily assessed. Additionally, it is conceivable that the monocolonization of humanized AhR mice with active strains of Bifidobacteria could further expand our knowledge on the health-promoting effect of these technologically interesting bacteria.

Interestingly, some of the identified Bifidobacteria for the AhR activator, were already described for their positive effects on intestinal homeostasis, such as *B. breve*. Yet, butyrate and butyrate-producing bacteria have been extensively described for their role in intestinal homeostasis and anti-inflammatory effects. Patient affected by inflammatory bowel diseases (IBD) showed a reduced abundance of butyrate-producing bacteria as well as of AhR ligands. Therefore it is tempting to speculate that the butyrate and Bifidobacteria-derived ligand, could positively impact the intestinal inflammation via AhR. Butyrate has been studied in clinical applications and some Bifidobacteria are already used as probiotics for their positive role on intestinal homeostasis. Thus, the identification of butyrate as a novel ligand and Bifidobacteria as activators of AhR could extend their therapeutic interest toward the treatment of inflammatory conditions via the AhR signalling, in the gut as well as in other anatomical sites.

Beyond the gut, AhR is described to be expressed in a variety of tissues, especially in other barrier organs so, it is conceivable that our described mechanisms could have a relevance in other anatomical sites beyond the digestive tract. AhR is expressed in the respiratory tract where AhR activation by TCDD in a lung epithelial cell line was shown to induced the expression of MUC5AC (Wong et al. 2010). Although, little is known about the presence of butyrate in the different sites of the respiratory tract, Actinobacteria have been identified as one of the dominant phyla in airways (Cabrera-Rubio et al. 2012). In particular, Bifidobacteria have been described in nasopharynx of children and adult, in anterior nares and oropharynx in elderly. It is therefore tempting to speculate that the described activation of AhR by Bifidobacteria could also occur in the airways. Bifidobacteria have also been identified as members of "healthy" vaginal microbiota (Freitas & Hill 2017) together with a variety of butyrate-producing species, associated with bacterial asymptomatic vaginosis (Aldunate et al.

2015). In this site AhR expression was detected in squamous epithelial cells (The Human Protein Atlas). As well, considering the expression of AhR in the skin and the complex microbiota, we could not exclude that butyrate and Bifidobacteria may be effective also in this anatomical site, although extremely difficult to study.

Consequently, our results for Bifidobacteria- and butyrate-induced activation of AhR in IECs, might be relevant in other anatomical sites beyond the gut, such as the airways and the vagina, in which Bifidobacteria and SCFA-producing bacteria (in vagina) found their ecological niche and could influence the physiology of the host via AhR.

Chapter 5. Supplementary

5.1. *Choline-TMA catabolism by gut microbiota*

(The following experiments have been performed in collaboration with Dr Dumas Marc-Emmanuel and Dr Chilloux Julien, Imperial College of London, London, UK)

5.1.1. Comments on Objectives and Methods

Included in the European project MetaCardis, part of the original PhD project aimed to identify new bacterial species and genes involved in choline-TMA catabolism, within the human gut microbiota. The interest in this metabolism arises from the evidenced that the microbial-derived TMA is converted, by hepatic enzymes, in the proatherogenic compound TMAO. Considering the strict microbial origin of TMA mainly derived from dietary choline and carnitine, the choline-TMA catabolism provides a link between microbiota, diet and cardiovascular diseases.

Through an *in silico* approach, different authors suggested that several members of the human intestinal microbiota (including *Clostridium* spp., *Anaerococcus* spp., *Collinsella* spp., *Desulfitobacterium* spp., *Klebsiella* spp., *Escherichia* spp., *Providencia* spp., *Yokenella* spp. and *Proteus* spp.) express the TMA-lyase and consequently are capable of degrading choline to TMA. However, Romano and colleagues reported the choline degradation and TMA production *in vivo* by a bacterial strain, *Edwardsiella tarda* ATCC23685, not to be predicted to code for TMA-lyase.

Our aim was to identify gut microbial genes involved in choline-TMA catabolism. To achieve this goal we aimed to screen (meta)genomic clones (*Escherichia coli* bearing fosmids with metagenomic inserts of ~40 kb in length and randomly selected from a global library of ~20000 clones) generated from a fecal sample of an healthy human or TMA-producing bacteria. However, prior to the screening, validations of the approach were performed by testing the host strain of the metagenomic library for its ability to growth in presence of TMA and for validating its inefficiency to produce or degrade one or both metabolites. Additionally validations were performed on the capacity of *Edwardsiella tarda* to degrade choline and produce TMA, although not predicted to have the cut gene cluster in its genome.

5.1.2. Experimental set-up

Cultures of *E. coli* EPI300 PCC1 were performed in presence of choline (60mM, 30mM, 15mM, 7.5mM, 3.75mM, 1mM, 0mM) and/or TMA (1µM, 0.1µM, 0.01µM, 0µM), grown in LB medium at 37°C.

In our experimental set-up, we cultured two described TMA-producers bacteria, *Clostridium asparagiforme* (DSMZ15981) and *Clostridium hathewayii* (DSMZ13479), and two strains of Edwardsiella *tarda* (DSM30052 and ATCC23685). The cultures of *E. tarda*, *C. asparagiforme* and *C. hathewayii* were supplemented in choline (0mM, 1mM and 30mM) to assess choline degradation and TMA production. Aliquots of these bacterial cultures were taken at 24h, 48h and 96h for each condition. Non-inoculated media was used as background controls.

The detection of choline degradation and TMA production was performed by standard ^{1}H NMR spectra measurement at Imperial College of London (ICL) on a spectrometer (Bruker, Rheinstetten, Germany) operating at 600.22 MHz ^{1}H frequency for all the bacterial samples as well as all the control media. The ^{1}H NMR spectra were phased, the baseline corrected and were imported into Matlab R2014a and normalized to the internal standard (trimethylsilyl propanoic acid). These evaluations were manually confirmed by detailed analysis of individual spectra using Chenomix Profiler 8.1.

5.1.3. Comments on the Results and Discussion

By ^{1}H-NMR analysis, we confirmed that *C. asparagiforme* and *C. hathewayii* are able to completely convert choline in TMA after 24h, as previously published, but we did not detect any choline consumption by the two strains of *E. tarda*.

Choline and TMA at different concentrations (up to 60mM of choline and 1µM of TMA) did not affect the growth of *E. coli* EPI300 PCC1 in supplemented LB medium during 24h of incubation (data not shown). However, control dilutions of TMA (up to 0.1µM) and choline (up to 60mM) in non-inoculated LB, revealed that the experimental approach by ^{1}H-NMR analysis was not sensible enough for the detection, especially for TMA, even in a simple non-inoculated dilution.

By expressing in *E. coli* the cut gene cluster of a strong TMA-producer bacteria, coherent with our (meta)genomic approach, Craciun showed a significant decrease in the TMA produced (Craciun & Balskus 2012). From about 500uM of TMA produced by the wild-type strain, they reported around 1.5uM of production when the entire gene cluster is expressed by *E. coli* (Craciun & Balskus 2012). Thus, considering this results, it is conceivable that our approach is not enough sensible and not adapted for the screening of a (meta)genomic library in which the expressed insert could derived from weak producers and the potential gene cluster could be fragmented. To get rid of this detection problem, other detection techniques should have been tested in order to ensure a correct detection of TMA in small volume and at low concentrations, as for example the uHPLC-MS/MS, as proposed by other authors (Romano 2015). Considering that up to day only *Edwardsiella tarda* is described to produced TMA independently of the cut gene, and in light of our negative results on its choline metabolization, it became questionable the existence of another gene cluster responsible for this metabolism. Based on these experimental results, we decided to cease this project and focus our attention on the work concerning AhR, which became the object of my thesis.

Bibliography

Abbott, B.D., Birnbaum, L.S. & Perdew, G.H., 1995. Developmental expression of two members of a new class of transcription factors: I. Expression of aryl hydrocarbon receptor in the C57BL/6N mouse embryo. *Developmental Dynamics*, 204(2), pp.133–143.

Abnet, C.C. et al., 1999. Two forms of aryl hydrocarbon receptor type 2 in rainbow trout (Oncorhynchus mykiss). Evidence for differential expression and enhancer specificity. *The Journal of biological chemistry*, 274(21), pp.15159–66.

Adachi, J. et al., 2001. Indirubin and Indigo are Potent Aryl Hydrocarbon Receptor Ligands Present in Human Urine. *Journal of Biological Chemistry*, 276(34), pp.31475–31478.

Al-Waiz, M. et al., 1992. The exogenous origin of trimethylamine in the mouse. *Metabolism*, 41(2), pp.135–136.

Aldunate, M. et al., 2015. Antimicrobial and immune modulatory effects of lactic acid and short chain fatty acids produced by vaginal microbiota associated with eubiosis and bacterial vaginosis. *Frontiers in physiology*, 6, p.1.

Alex, S. et al., 2013. Short-Chain Fatty Acids Stimulate Angiopoietin-Like 4 Synthesis in Human Colon Adenocarcinoma Cells by Activating Peroxisome Proliferator-Activated Receptor . *Molecular and Cellular Biology*, 33(7), pp.1303–1316.

Alexeev, E.E. et al., 2018. Microbiota-Derived Indole Metabolites Promote Human and Murine Intestinal Homeostasis through Regulation of Interleukin-10 Receptor. *The American Journal of Pathology*.

Amar, J. et al., 2011. Intestinal mucosal adherence and translocation of commensal bacteria at the early onset of type 2 diabetes: molecular mechanisms and probiotic treatment. *EMBO Molecular Medicine*, 3(9), pp.559–572.

Andersson, P. et al., 2002. A constitutively active dioxin/aryl hydrocarbon receptor induces stomach tumors. *Proceedings of the National Academy of Sciences of the United States of America*, 99(15), pp.9990–5.

Antonsson, C. et al., 1995. Distinct roles of the molecular chaperone hsp90 in modulating dioxin receptor function via the basic helix-loop-helix and PAS domains. *Molecular and cellular biology*, 15(2), pp.756–65.

Aron-Wisnewsky, J. & Clément, K., 2016. The gut microbiome, diet, and links to cardiometabolic and chronic disorders. *Nature Reviews Nephrology*, 12(3), pp.169–181.

Arumugam, M. et al., 2011. Enterotypes of the human gut microbiome. *Nature*, 473(7346), pp.174–180.

Atuma, C. et al., 2001. The adherent gastrointestinal mucus gel layer: thickness and physical state in vivo. *American Journal of Physiology-Gastrointestinal and Liver Physiology*, 280(5), pp.G922–G929.

Baba, T. et al., 2001. Structure and Expression of the Ah Receptor Repressor Gene. *Journal of Biological Chemistry*, 276(35), pp.33101–33110.

Bach, J.-F., 2002. The Effect of Infections on Susceptibility to Autoimmune and Allergic Diseases. *New England Journal of Medicine*, 347(12), pp.911–920.

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.

Backlund, M. & Ingelman-Sundberg, M., 2005. Regulation of aryl hydrocarbon receptor signal transduction by protein tyrosine kinases. *Cellular Signalling*, 17(1), pp.39–48.

Bacsi, S.G., Reisz-Porszasz, S. & Hankinson, O., 1995. Orientation of the heterodimeric aryl hydrocarbon (dioxin) receptor complex on its asymmetric DNA recognition sequence. *Molecular Pharmacology*, 47(3),

pp.432–438.

Baeuerle, P.A. & Baichwal, V.R., 1997. NF-kappa B as a frequent target for immunosuppressive and anti-inflammatory molecules. *Advances in immunology*, 65, pp.111–37.

Bansal, T. et al., 2010. The bacterial signal indole increases epithelial-cell tight-junction resistance and attenuates indicators of inflammation. *Proceedings of the National Academy of Sciences*, 107(1), pp.228–233.

Barnaba, C. et al., 2017. The catalytic function of cytochrome P450 is entwined with its membrane-bound nature. *F1000Research*, 6, p.662.

Barouki, R., Coumoul, X. & Fernandez-Salguero, P.M., 2007. The aryl hydrocarbon receptor, more than a xenobiotic-interacting protein. *FEBS Letters*, 581(19), pp.3608–3615.

Bashir, M.E.H. et al., 2004. Toll-like receptor 4 signaling by intestinal microbes influences susceptibility to food allergy. *Journal of immunology (Baltimore, Md. : 1950)*, 172(11), pp.6978–87.

Batterham, R.L. & Bloom, S.R., 2003. The Gut Hormone Peptide YY Regulates Appetite. *Annals of the New York Academy of Sciences*, 994(1), pp.162–168.

Beischlag, T. V. & Perdew, G.H., 2005. ERa-AHR-ARNT protein-protein interactions mediate estradiol-dependent transrepression of dioxin-inducible gene transcription. *Journal of Biological Chemistry*, 280(22), pp.21607–21611.

Bendheim, P.E. et al., 2002. Development of indole-3-propionic acid (OXIGON™) for alzheimer's disease. *Journal of Molecular Neuroscience*, 19(1–2), pp.213–217.

Benson, J.M. & Shepherd, D.M., 2011. Aryl hydrocarbon receptor activation by TCDD reduces inflammation associated with Crohn's disease. *Toxicological sciences : an official journal of the Society of Toxicology*, 120(1), pp.68–78.

Bergander, L. et al., 2004. Metabolic fate of the Ah receptor ligand 6-formylindolo[3,2-b]carbazole. *Chemico-Biological Interactions*, 149(2–3), pp.151–164.

Bernink, J.H. et al., 2013. Human type 1 innate lymphoid cells accumulate in inflamed mucosal tissues. *Nature Immunology*, 14(3), pp.221–229.

Biagi, E. et al., 2010. Through Ageing, and Beyond: Gut Microbiota and Inflammatory Status in Seniors and Centenarians N. Ahmed, ed. *PLoS ONE*, 5(5), p.e10667.

Bisson, W.H. et al., 2009. Modeling of the aryl hydrocarbon receptor (AhR) ligand binding domain and its utility in virtual ligand screening to predict new AhR ligands. *Journal of Medicinal Chemistry*, 52(18), pp.5635–5641.

Bjeldanes, L.F. et al., 1991. Aromatic hydrocarbon responsiveness-receptor agonists generated from indole-3-carbinol in vitro and in vivo: comparisons with 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Proceedings of the National Academy of Sciences of the United States of America*, 88(21), pp.9543–9547.

Bjursell, M. et al., 2011. Improved glucose control and reduced body fat mass in free fatty acid receptor 2-deficient mice fed a high-fat diet. *American Journal of Physiology-Endocrinology and Metabolism*, 300(1), pp.E211–E220.

Browne, H.P. et al., 2016. Culturing of "unculturable" human microbiota reveals novel taxa and extensive sporulation. *Nature*, 533(7604), pp.543–546.

Buonocore, S. et al., 2010. Innate lymphoid cells drive interleukin-23-dependent innate intestinal pathology. *Nature*, 464(7293), pp.1371–1375.

Butler, R.A. et al., 2001. An aryl hydrocarbon receptor (AHR) homologue from the soft-shell clam, Mya arenaria: evidence that invertebrate AHR homologues lack 2,3,7,8-tetrachlorodibenzo-p-dioxin and β-

naphthoflavone binding. *Gene*, 278(1–2), pp.223–234.

Cabrera-Rubio, R. et al., 2012. Microbiome diversity in the bronchial tracts of patients with chronic obstructive pulmonary disease. *Journal of clinical microbiology*, 50(11), pp.3562–8.

Camelo, A. et al., 2017. IL-33, IL-25, and TSLP induce a distinct phenotypic and activation profile in human type 2 innate lymphoid cells. *Blood Advances*, 1(10), pp.577–589.

Cantarel, B.L. et al., 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research*, 37(Database), pp.D233–D238.

Cario, E., Gerken, G. & Podolsky, D.K., 2004. Toll-like receptor 2 enhances ZO-1-associated intestinal epithelial barrier integrity via protein kinase C. *Gastroenterology*, 127(1), pp.224–38.

Carver, L.A. & Bradfield, C.A., 1997. Ligand-dependent interaction of the aryl hydrocarbon receptor with a novel immunophilin homolog in vivo. *Journal of Biological Chemistry*, 272(17), pp.11452–6.

Carver, L.A., Jackiw, V. & Bradfields, C.A., 1994. The 90-kDa Heat Shock Protein Is Essential for Ah Receptor Signaling in a Yeast Expression System. *THE JOURNAL OF BIOLOGICAL CHEMISTRY*, 269(48), pp.30109–30112.

Chang, C.Y. & Puga, A., 1998. Constitutive activation of the aromatic hydrocarbon receptor. *Molecular and cellular biology*, 18(1), pp.525–35.

Chapman-Smith, A., Lutwyche, J.K. & Whitelaw, M.L., 2004. Contribution of the Per/Arnt/Sim (PAS) Domains to DNA Binding by the Basic Helix-Loop-Helix PAS Transcriptional Regulators. *Journal of Biological Chemistry*, 279(7), pp.5353–5362.

Le Chatelier, E. et al., 2013. Richness of human gut microbiome correlates with metabolic markers. *Nature*, 500(7464), pp.541–546.

Chen, H.S., Singh, S.S. & Perdew, G.H., 1997. The Ah receptor is a sensitive target of geldanamycin-induced protein turnover. *Archives of Biochemistry and Biophysics*, 348(1), pp.190–198. A

Chen, X. et al., 2011. TGR5: a novel target for weight maintenance and glucose metabolism. *Experimental diabetes research*, 2011, p.853501.

Chen, Y. et al., 2007. Lysine propionylation and butyrylation are novel post-translational modifications in histones. *Molecular & cellular proteomics : MCP*, 6(5), pp.812–9.

Chimerel, C. et al., 2014. Bacterial metabolite indole modulates incretin secretion from intestinal enteroendocrine L cells. *Cell reports*, 9(4), pp.1202–8.

Chinen, I. et al., 2015. The aryl hydrocarbon receptor/microRNA-212/132 axis in T cells regulates IL-10 production to maintain intestinal homeostasis. *International Immunology*, 27(8), pp.405–415.

Cohen, L.J. et al., 2015. Functional metagenomic discovery of bacterial effectors in the human microbiome and isolation of commendamide, a GPCR G2A/132 agonist. *Proceedings of the National Academy of Sciences*, 112(35), pp.E4825–E4834.

Collado, M.C. et al., 2016. Human gut colonisation may be initiated in utero by distinct microbial communities in the placenta and amniotic fluid. *Scientific Reports*, 6(1), p.23129.

Costea, P.I. et al., 2017. Enterotypes in the landscape of gut microbial community composition. *Nature Microbiology*, 3(1), pp.8–16.

Cotillard, A. et al., 2013. Dietary intervention impact on gut microbial gene richness. *Nature*, 500(7464), pp.585–588.

Coumailleau, P. et al., 1995. Definition of a minimal domain of the dioxin receptor that is associated with Hsp90 and maintains wild type ligand binding affinity and specificity. *Journal of Biological Chemistry*,

270(42), pp.25291–25300.

Couturier-Maillard, A. et al., 2013. NOD2-mediated dysbiosis predisposes mice to transmissible colitis and colorectal cancer. *Journal of Clinical Investigation*, 123(2), pp.700–11.

Cox, M.B., Miller, C.A. & III, 2004. Cooperation of heat shock protein 90 and p23 in aryl hydrocarbon receptor signaling. *Cell stress & chaperones*, 9(1), pp.4–20.

Craciun, S. & Balskus, E.P., 2012. Microbial conversion of choline to trimethylamine requires a glycyl radical enzyme. *PNAS*, 109, pp.21307–21312.

Cummings, J.H. et al., 1987. Short chain fatty acids in human large intestine, portal, hepatic and venous blood. *Gut*, 28(10), pp.1221–7.

Daly, K. & Shirazi-Beechey, S.P., 2006. Microarray analysis of butyrate regulated genes in colonic epithelial cells. *DNA and cell biology*, 25(1), pp.49–62.

Darfeuille-Michaud, A., 2002. Adherent-invasive Escherichia coli: a putative new E. coli pathotype associated with Crohn's disease. *International Journal of Medical Microbiology*, 292(3–4), pp.185–193.

De Vadder, F. et al., 2014. Microbiota-Generated Metabolites Promote Metabolic Benefits via Gut-Brain Neural Circuits. *Cell*, 156(1–2), pp.84–96.

Denison, M.S. & Nagy, S.R., 2003. Activation of the aryl hydrocarbon receptor by structurally diverse exogenous and endogenous chemicals. *Annual review of pharmacology and toxicology*, 43, pp.309–334.

Derrien, M. et al., 2008. The Mucin degrader Akkermansia muciniphila is an abundant resident of the human intestinal tract. *Applied and environmental microbiology*, 74(5), pp.1646–8.

Desreumaux, P. et al., 2001. Attenuation of colon inflammation through activators of the retinoid X receptor (RXR)/peroxisome proliferator-activated receptor gamma (PPARgamma) heterodimer. A basis for new therapeutic strategies. *The Journal of experimental medicine*, 193(7), pp.827–838.

Dethlefsen, L., McFall-Ngai, M. & Relman, D.A., 2007. An ecological and evolutionary perspective on human–microbe mutualism and disease. *Nature*, 449(7164), pp.811–818.

Dethlefsen, L. & Relman, D.A., 2011. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proceedings of the National Academy of Sciences*, 108(Supplement_1), pp.4554–4561.

DiMeglio, P. et al., 2014. Activation of the aryl hydrocarbon receptor dampens the severity of inflammatory skin conditions. *Immunity*, 40, pp.989–1001.

Ding, L. et al., 2003. Ligand-independent activation of estrogen receptor alpha by XBP-1. *Nucleic acids research*, 31(18), pp.5266–74.

Donaldson, G.P., Lee, S.M. & Mazmanian, S.K., 2016. Gut biogeography of the bacterial microbiota. *Nature reviews. Microbiology*, 14(1), pp.20–32.

Donnelly, D., 2012. The structure and function of the glucagon-like peptide-1 receptor and its ligands. *British Journal of Pharmacology*, 166(1), pp.27–41.

El-Tanani, M.K.K. & Green, C.D., 1997. Two Separate Mechanisms for Ligand-Independent Activation of the Estrogen Receptor. *Molecular Endocrinology*, 11(7), pp.928–937.

Ema, M. et al., 1994. Dioxin binding activities of polymorphic forms of mouse and human aryl hydrocarbon receptors. *J Biol.Chem*, 269(44), pp.27337–27343.

Esser, C. & Rannug, A., 2015. The Aryl Hydrocarbon Receptor in Barrier Organ Physiology, Immunology, and Toxicology. *Pharmacological reviews*, 67(2), pp.259–279.

Evans, B.R. et al., 2008. Repression of aryl hydrocarbon receptor (AHR) signaling by AHR repressor: role of DNA binding and competition for AHR nuclear translocator. *Molecular Pharmacology*, 73(2), pp.387–398.

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.9066–9071.

Fellows, R. et al., 2018. Microbiota derived short chain fatty acids promote histone crotonylation in the colon through histone deacetylases. *Nature Communications*, 9(1), p.105.

Ferguson, A., 1977. Intraepithelial lymphocytes of the small intestine. *Gut*, 18(11), pp.921–37.

Fernandez-Salguero, P.M. et al., 1995. Immune System Impairment and Hepatic Fibrosis in Mice Lacking the Dioxin- Binding Ah Receptor. *Science*, 268, pp.722–726.

Fernandez-Salguero, P.M. et al., 1997. Lesions of Aryl-hydrocarbon Receptor-deficient Mice. *Vet Pathol*, 34, pp.605–614.

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 of the United States of America*, 107(33), pp.14691–6.

Fitzgerald, C.T., Nebert, D.W. & Puga, A., 1998. Regulation of mouse Ah receptor (Ahr) gene basal expression by members of the Sp family of transcription factors. *DNA and cell biology*, 17(9), pp.811–22.

Flaveny, C.A. et al., 2009. Ligand Selectivity and Gene Regulation by the Human Aryl Hydrocarbon Receptor in Transgenic Mice. *Molecular Pharmacology*, 75(6), pp.1412–1420.

Flint, H.J. et al., 2008. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nature reviews. Microbiology*, 6(2), pp.121–31.

Franzosa, E.A. et al., 2015. Identifying personal microbiomes using metagenomic codes. *Proceedings of the National Academy of Sciences of the United States of America*, 112(22), pp.E2930-8.

Freitas, A.C. & Hill, J.E., 2017. Quantification, isolation and characterization of Bifidobacterium from the vaginal microbiomes of reproductive aged women. *Anaerobe*, 47, pp.145–156.

Frericks, M., Meissner, M. & Esser, C., 2007. Microarray analysis of the AHR system: Tissue-specific flexibility in signal and target genes. *Toxicology and Applied Pharmacology*, 220(3), pp.320–332.

Fribourgh, J.L. & Partch, C.L., 2017. Assembly and function of bHLH–PAS complexes. *PNAS*, 114(21), pp.5330–5332.

Fujii-Kuriyama, Y. et al., 1994. Ah receptor: a novel ligand-activated transcription factor. *Experimental and clinical immunogenetics*, 11(2–3), pp.65–74.

Fukumoto, S. et al., 2014. Identification of a probiotic bacteria-derived activator of the aryl hydrocarbon receptor that inhibits colitis. *Immunology and cell biology*, 92(5), pp.460–5.

Furumatsu, K. et al., 2011. A Role of the Aryl Hydrocarbon Receptor in Attenuation of Colitis. *Digestive Diseases and Sciences*, 56(9), pp.2532–2544.

Gao, J. et al., 2018. Impact of the Gut Microbiota on Intestinal Immunity Mediated by Tryptophan Metabolism. *Frontiers in Cellular and Infection Microbiology*, 8, p.13.

Garcia, M.A., Nelson, W.J. & Chavez, N., 2017. Cell–Cell Junctions Organize Structural and Signaling Networks. *Cold Spring Harbor Perspectives in Biology*, 10(4), p.a029181.

Gavini, F. et al., 2001. Differences in the Distribution of Bifidobacterial and Enterobacterial Species in Human Faecal Microflora of Three Different (Children, Adults, Elderly) Age Groups. *Microbial Ecology*

in Health and Disease, 13(1), pp.40–45.

Geirnaert, A. et al., 2017. Butyrate-producing bacteria supplemented in vitro to Crohn's disease patient microbiota increased butyrate production and enhanced intestinal epithelial barrier integrity. *Scientific Reports*, 7(1), p.11450.

Gerbal-Chaloin, S. et al., 2006. Role of CYP3A4 in the regulation of the aryl hydrocarbon receptor by omeprazole sulphide. *Cellular Signalling*, 18(5), pp.740–750.

Gerbe, F. et al., 2011. Distinct ATOH1 and Neurog3 requirements define tuft cells as a new secretory cell type in the intestinal epithelium. *The Journal of cell biology*, 192(5), pp.767–80.

Gerbe, F. & Jay, P., 2016. Intestinal tuft cells: Epithelial sentinels linking luminal cues to the immune system. *Mucosal Immunology*, 9(6), pp.1353–1359.

Geremia, A. et al., 2011. IL-23–responsive innate lymphoid cells are increased in inflammatory bowel disease. *The Journal of Experimental Medicine*, 208(6), pp.1127–1133.

Ghotbaddini, M. & Powell, J.B., 2015. The AhR Ligand, TCDD, Regulates Androgen Receptor Activity Differently in Androgen-Sensitive versus Castration-Resistant Human Prostate Cancer Cells. *International Journal of Environmental Research and Public Health*, 12(7), pp.7506–18.

Gillam, E.M.J. et al., 1999. Formation of indigo by recombinant mammalian cytochrome P450. *Biochemical and Biophysical Research Communications*, 265(2), pp.469–472.

Girardin, S.E., Boneca, I.G., et al., 2003. Nod1 detects a unique muropeptide from gram-negative bacterial peptidoglycan. *Science*, 300(5625), pp.1584–1587.

Girardin, S.E., Boneca, I.G., et al., 2003. Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection. *Journal of Biological Chemistry*, 278(11), pp.8869–8872.

Goudot, C. et al., 2017. Aryl Hydrocarbon Receptor Controls Monocyte Differentiation into Dendritic Cells versus Macrophages. *Immunity*, 47(3), p.582–596.e6.

Grönlund, M.-M. et al., 2007. Maternal breast-milk and intestinal bifidobacteria guide the compositional development of the Bifidobacterium microbiota in infants at risk of allergic disease. *Clinical & Experimental Allergy*, 37(12), pp.1764–1772.

Guyonnet Duperat, V. et al., 1995. Characterization of the human mucin gene MUC5AC: a consensus cysteine-rich domain for 11p15 mucin genes? *The Biochemical journal*, 305 (Pt 1)(Pt 1), pp.211–9.

Ha, C.W., Lam, Y.Y. & Holmes, A.J., 2014. Mechanistic links between gut microbial community dynamics, microbial functions and metabolic health. *World Journal of Gastroenterology*, 20(44), pp.16498–16517.

Haarmann-Stemmann, T. et al., 2007. Analysis of the transcriptional regulation and molecular function of the aryl hydrocarbon receptor repressor in human cell lines. *Drug metabolism and disposition: the biological fate of chemicals*, 35(12), pp.2262–2269.

Haarmann-Stemmann, T., Bothe, H. & Abel, J., 2009. Growth factors, cytokines and their receptors as downstream targets of arylhydrocarbon receptor (AhR) signaling pathways. *Biochemical Pharmacology*, 77(4), pp.508–520.

Haas, K. et al., 2016. Aryl Hydrocarbon Receptor in Keratinocytes Is Essential for Murine Skin Barrier Integrity. *Journal of Investigative Dermatology*, 136(11), pp.2260–2269.

Hahn, M.E., 2002. Aryl hydrocarbon receptors: Diversity and evolution. *Chemico-Biological Interactions*, 141(1–2), pp.131–160.

Hahn, M.E., Allan, L.L. & Sherr, D.H., 2009. *Regulation of constitutive and inducible AHR signaling: Complex interactions involving the AHR repressor*,

Halim, T.Y.F. et al., 2012. Lung Natural Helper Cells Are a Critical Source of Th2 Cell-Type Cytokines in Protease Allergen-Induced Airway Inflammation. *Immunity*, 36(3), pp.451–463..

Halme, L. et al., 2006. Family and twin studies in inflammatory bowel disease. *World journal of gastroenterology*, 12(23), pp.3668–72.

Hamada, H. et al., 2002. Identification of multiple isolated lymphoid follicles on the antimesenteric wall of the mouse small intestine. *Journal of immunology (Baltimore, Md. : 1950)*, 168(1), pp.57–64.

Han, B. et al., 2016. Aryl Hydrocarbon Receptor Activation in Intestinal Obstruction Ameliorates Intestinal Barrier Dysfunction Via Suppression of MLCK-MLC Phosphorylation Pathway. *SHOCK*, 46(3), pp.319–328.

Han, E.H. et al., 2015. Endosulfan induces cyp1a1 expression mediated through aryl hydrocarbon receptor signal transduction by protein kinase c. *Toxicological Research*, 31(4), pp.339–345.

Hao, N. et al., 2011. Identification of residues in the N-terminal PAS domains important for dimerization of Arnt and AhR. *Nucleic Acids Research*, 39(9), pp.3695–3709.

Harper, P.A., Golas, C.L. & Okey, A.B., 1988. Characterization of the Ah Receptor and Aryl Hydrocarbon Hydroxylase Induction by 2,3,7,8-Tetrachlorodibenzo-p-dioxin and Benz(a)anthracene in the Human A431 Squamous Cell Carcinoma Line. *Cancer Research*, 48(9), pp.2388–2395.

Hashimoto, T. et al., 2012. ACE2 links amino acid malnutrition to microbial ecology and intestinal inflammation. *Nature*, 487, pp.477–483.

Heath-Pagliuso, S. et al., 1998. Activation of the Ah receptor by tryptophan and tryptophan metabolites. *Biochemistry*, 37(33), pp.11508–11515.

Heid, S.E., Pollenz, R.S. & Swanson, H.I., 2000. Role of heat shock protein 90 dissociation in mediating agonist-induced activation of the aryl hydrocarbon receptor. *Mol Pharmacol*, 57(1), pp.82–92.

Henry, E.C. et al., 2006. A potential endogenous ligand for the aryl hydrocarbon receptor has potent agonist activity in vitro and in vivo. *Archives of Biochemistry and Biophysics*, 450(1), pp.67–77.

Higashimura, Y. et al., 2015. Propionate promotes fatty acid oxidation through the up-regulation of peroxisome proliferator-activated receptor α in intestinal epithelial cells. *Journal of Nutritional Science and Vitaminology*, 61(6), pp.511–515.

Hippe, B. et al., 2016. *Faecalibacterium prausnitzii* phylotypes in type two diabetic, obese, and lean control subjects. *Beneficial Microbes*, 7(4), pp.511–517.

Holzer, P., Reichmann, F. & Farzi, A., 2012. Neuropeptide Y, peptide YY and pancreatic polypeptide in the gut–brain axis. *Neuropeptides*, 46(6), pp.261–274.

Hooper, L. V., 2011. You AhR what you eat: linking diet and immunity. *Cell*, 147(3), pp.489–91.

Hooper, L. V & Macpherson, A.J., 2010. Immune adaptations that maintain homeostasis with the intestinal microbiota. *Nat Rev Immunol*, 10(3), pp.159–169.

Hosoya, T. et al., 2008. Inducibility of cytochrome P450 1A1 and chemical carcinogenesis by benzo[a]pyrene in AhR repressor-deficient mice. *Biochemical and Biophysical Research Communications*, 365(3), pp.562–567.

Hou, C. et al., 2015. Intestinal microbiota succession and immunomodulatory consequences after introduction of lactobacillus reuteri I5007 in neonatal piglets G. Loh, ed. *PLoS ONE*, 10(3), p.e0119505.

Hsieh, C.-Y. et al., 2015. Strengthening of the intestinal epithelial tight junction by *Bifidobacterium bifidum*. *Physiological Reports*, 3(3), p.e12327.

Huang, M.H., Lee, S.T. & Rajendran, K., 1997. A fresh cadaveric study of the paratubal muscles:

implications for eustachian tube function in cleft palate. *Plastic and reconstructive surgery*, 100(4), pp.833–42.

Huang, X., Powell-Coffman, J.A. & Jin, Y., 2004. The AHR-1 aryl hydrocarbon receptor and its co-factor the AHA-1 aryl hydrocarbon receptor nuclear translocator specify GABAergic neuron cell fate in C. elegans. *Development*, 131(4), pp.819–828.

Hubbard, T.D. et al., 2015. Adaptation of the human aryl hydrocarbon receptor to sense microbiota-derived indoles. *Scientific Reports*, 5(1), p.12689.

Ikuta, T. et al., 1998. Nuclear localization and export signals of the human aryl hydrocarbon receptor. *The Journal of biological chemistry*, 273(5), pp.2895–2904.

Ikuta, T. et al., 2000. Nucleocytoplasmic shuttling of the aryl hydrocarbon receptor. *Journal of biochemistry*, 127(3), pp.503–509.

Ikuta, T., Kobayashi, Y. & Kawajiri, K., 2004. Phosphorylation of nuclear localization signal inhibits the ligand-dependent nuclear import of aryl hydrocarbon receptor. *Biochemical and Biophysical Research Communications*, 317(2), pp.545–550.

Inagaki, T. et al., 2006. Regulation of antibacterial defense in the small intestine by the nuclear bile acid receptor. *Proceedings of the National Academy of Sciences*, 103(10), pp.3920–3925.

Ismail, A.S. et al., 2011. intraepithelial lymphocytes are essential mediators of host-microbial homeostasis at the intestinal mucosal surface. *Proceedings of the National Academy of Sciences*, 108(21), pp.8743–8748.

Isomäki, A.M., 1973. A new cell type (tuft cell) in the gastrointestinal mucosa of the rat. A transmission and scanning electron microscopic study. *Acta pathologica et microbiologica Scandinavica. Section A, Pathology*, p.Suppl 240:1-35.

Ito, T. et al., 2004. A Constitutively Active Arylhydrocarbon Receptor Induces Growth Inhibition of Jurkat T Cells through Changes in the Expression of Genes Related to Apoptosis and Cell Cycle Arrest. *Journal of Biological Chemistry*, 279(24), pp.25204–25210.

Iwabuchi, N. et al., 2007. In vitro Th1 cytokine-independent Th2 suppressive effects of bifidobacteria. *Microbiology and immunology*, 51(7), pp.649–60.

Iwanaga, T. et al., 2006. Cellular expression of monocarboxylate transporters (MCT) in the digestive tract of the mouse, rat, and humans, with special reference to slc5a8. *Biomedical research (Tokyo, Japan)*, 27(5), pp.243–54.

Jeon, S.G. et al., 2012. Probiotic Bifidobacterium breve Induces IL-10-Producing Tr1 Cells in the Colon. *PLoS Pathogens*, 8(5), p.e1002714.

Jin, U.-H. et al., 2014. Microbiome-Derived Tryptophan Metabolites and Their Aryl Hydrocarbon Receptor-Dependent Agonist and Antagonist Activities. *Molecular Pharmacology*, 85(May), pp.777–788.

Jin, U.-H. et al., 2017. Short Chain Fatty Acids Enhance Aryl Hydrocarbon (Ah) Responsiveness in Mouse Colonocytes and Caco-2 Human Colon Cancer Cells. *Scientific Reports*, 7(1), p.10163.

Jinno, A. et al., 2006. Induction of cytochrome P450-1A by the equine estrogen equilenin, a new endogenous aryl hydrocarbon receptor ligand. *Journal of Steroid Biochemistry and Molecular Biology*, 98(1), pp.48–55.

Johansson, M.E. V. et al., 2008. The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proceedings of the National Academy of Science*, 105(39), pp.15064–9.

Jones, B. V. et al., 2008. Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome. *Proceedings of the National Academy of Sciences*, 105(36), pp.13580–13585.

Jones, S., 2004. An overview of the basic helix-loop-helix proteins. *Genome Biology*, 5(6), p.226.

Jostins, L. et al., 2012. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature*, 491(7422), pp.119–124.

Kakiyama, G. et al., 2014. A simple and accurate HPLC method for fecal bile acid profile in healthy and cirrhotic subjects: validation by GC-MS and LC-MS. *Journal of Lipid Research*, 55(5), pp.978–990.

Kandori, H. et al., 1996. Histochemical, lectin-histochemical and morphometrical characteristics of intestinal goblet cells of germfree and conventional mice. *Experimental animals*, 45(2), pp.155–60.

Kaser, A., Martínez-Naves, E. & Blumberg, R.S., 2010. Endoplasmic reticulum stress: implications for inflammatory bowel disease pathogenesis. *Current Opinion in Gastroenterology*, 26(4), pp.318–326.

Kashiwagi, I. et al., 2015. Smad2 and Smad3 Inversely Regulate TGF-β Autoinduction in Clostridium butyricum-Activated Dendritic Cells. *Immunity*, 43(1), pp.65–79.

Katiyar, S.K., Matsui, M.S. & Mukhtar, H., 2000. Ultraviolet-B exposure of human skin induces cytochromes P450 1A1 and 1B1. *Journal of Investigative Dermatology*, 114(2), pp.328–333.

Kawajiri, K. et al., 2009. Aryl hydrocarbon receptor suppresses intestinal carcinogenesis in ApcMin/+ mice with natural ligands. *Proceedings of the National Academy of Sciences*, 106(32), pp.13481–13486.

Kawajiri, K. & Fujii-Kuriyama, Y., 2007. Cytochrome P450 gene regulation and physiological functions mediated by the aryl hydrocarbon receptor. *Archives of Biochemistry and Biophysics*, 464(2), pp.207–212.

Kawasaki, T. & Kawai, T., 2014. Toll-like receptor signaling pathways. *Frontiers in immunology*, 5, p.461.

Kazlauskas, A., Poellinger, L. & Pongratz, I., 1999. Evidence that the co-chaperone p23 regulates ligand responsiveness of the dioxin (Aryl hydrocarbon) receptor. *Journal of Biological Chemistry*, 274(19), pp.13519–24.

Ke, S. et al., 2001. Mechanism of suppression of cytochrome P-450 1A1 expression by tumor necrosis factor-alpha and lipopolysaccharide. *The Journal of biological chemistry*, 276(43), pp.39638–44.

Kellingray, L. et al., 2017. Consumption of a diet rich in Brassica vegetables is associated with a reduced abundance of sulphate-reducing bacteria: A randomised crossover study. *Mol. Nutr. Food Reyes.*, 61(9), p.1600992.

Kerkvliet, N.I., 2009. AHR-mediated immunomodulation: The role of altered gene transcription. *Biochemical Pharmacology*, 77, pp.746–760.

Kewley, R.J., Whitelaw, M.L. & Chapman-Smith, A., 2004. The mammalian basic helix–loop–helix/PAS family of transcriptional regulators. *The International Journal of Biochemistry & Cell Biology*, 36(36), pp.189–204.

Khosravi, A. et al., 2014. Gut microbiota promote hematopoiesis to control bacterial infection. *Cell host & microbe*, 15(3), pp.374–81.

Kim, C.J. et al., 2010. l-Tryptophan exhibits therapeutic function in a porcine model of dextran sodium sulfate (DSS)-induced colitis. *The Journal of Nutritional Biochemistry*, 21(6), pp.468–475.

Kim, D.W. et al., 2000. The RelA NF-kappaB subunit and the aryl hydrocarbon receptor (AhR) cooperate to transactivate the c-myc promoter in mammary cells. *Oncogene*, 19(48), pp.5498–5506.

Kimura, A. et al., 2009. Aryl hydrocarbon receptor in combination with Stat1 regulates LPS-induced inflammatory responses. *Journal of Experimental Medicine*, 206(9), pp.2027–2035.

Kimura, A. et al., 2008. Aryl hydrocarbon receptor regulates Stat1 activation and participates in the development of Th17 cells. *Proceedings of the National Academy of Sciences*, 105(28), pp.9721–9726.

Kiss, E. a. et al., 2011. Natural Aryl Hydrocarbon Receptor Ligands Control Organogenesis of Intestinal Lymphoid Follicles. *Science*, 334(6062), pp.1561–1565.

de Kivit, S. et al., 2013. Intestinal Epithelium-Derived Galectin-9 Is Involved in the Immunomodulating Effects of Nondigestible Oligosaccharides. *Journal of Innate Immunity*, 5(6), pp.625–638.

Klaassens, E.S. et al., 2009. Mixed-Species Genomic Microarray Analysis of Fecal Samples Reveals Differential Transcriptional Responses of Bifidobacteria in Breast- and Formula-Fed Infants. *Applied and Environmental Microbiology*, 75(9), pp.2668–2676.

Ko, H.P. et al., 1997. Transactivation domains facilitate promoter occupancy for the dioxin-inducible CYP1A1 gene in vivo. *Molecular and cellular biology*, 17(7), pp.3497–3507.

Kobayashi, K.S. et al., 2005. Nod2-Dependent Regulation of Innate and Adaptive Immunity in the Intestinal Tract. *Advancement Of Science*, 307(5710), pp.731–734.

Koeth, R. a et al., 2013. Intestinal microbiota metabolism of l-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nature Medicine*, 19(5), pp.576–585.

Korkalainen, M., Tuomisto, J. & Pohjanvirta, R., 2001. The AH receptor of the most dioxin-sensitive species, guinea pig, is highly homologous to the human AH receptor. *Biochemical and biophysical research communications*, 285(5), pp.1121–9.

Koyasu, S. & Moro, K., 2012. Role of innate lymphocytes in infection and inammation. *Frontiers in Immunology*, 3(MAY), p.101.

Kumar, M.B. et al., 2001. The Q-rich Subdomain of the Human Ah Receptor Transactivation Domain Is Required for Dioxin-mediated Transcriptional Activity. *Journal of Biological Chemistry*, 276(45), pp.42302–42310.

Kurokawa, K. et al., 2007. Comparative Metagenomics Revealed Commonly Enriched Gene Sets in Human Gut Microbiomes. *DNA Research*, 14(4), pp.169–181.

Lagier, J.-C. et al., 2012. Microbial culturomics: paradigm shift in the human gut microbiome study. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, 18(12), pp.1185–93.

Lamas, B. et al., 2016. CARD9 impacts colitis by altering gut microbiota metabolism of tryptophan into aryl hydrocarbon receptor ligands. *Nature medicine*, 22(6), pp.598–605.

Lamas, B., Natividad, J.M. & Sokol, H., 2018. Aryl hydrocarbon receptor and intestinal immunity. *Mucosal Immunology*, (August 2017).

Larsson, E. et al., 2012. Analysis of gut microbial regulation of host gene expression along the length of the gut and regulation of gut microbial ecology through MyD88. *Gut*, 61(8), pp.1124–31.

Lee, A.J. et al., 2003. Characterization of the Oxidative Metabolites of 17β-Estradiol and Estrone Formed by 15 Selectively Expressed Human Cytochrome P450 Isoforms. *Endocrinology*, 144(8), pp.3382–3398.

Lee, J. et al., 2006. Maintenance of colonic homeostasis by distinctive apical TLR9 signalling in intestinal epithelial cells. *Nature Cell Biology*, 8(12), pp.1327–1336.

Lee, J.H., Wood, T.K. & Lee, J., 2015. Roles of indole as an interspecies and interkingdom signaling molecule. *Trends in Microbiology*, 23(11), pp.707–718.

Lee, J.S. et al., 2012. AHR drives the development of gut ILC22 cells and postnatal lymphoid tissues via pathways dependent on and independent of Notch. *Nat Immunol*, 13(2), pp.144–151.

Lees, M.J., Peet, D.J. & Whitelaw, M.L., 2003. Defining the role for XAP2 in stabilization of the dioxin receptor. *Journal of Biological Chemistry*, 278(38), pp.35878–35888.

Lepage, P. et al., 2011. Twin study indicates loss of interaction between microbiota and mucosa of patients with ulcerative colitis. *Gastroenterology*, 141(1), pp.227–236.

Ley, R.E. et al., 2005. Obesity alters gut microbial ecology. *Proceedings of the National Academy of Sciences of the United States of America*, 102(31), pp.11070–5.

Li, H. et al., 2015. The outer mucus layer hosts a distinct intestinal microbial niche. *Nature Communications*, 6(1), p.8292.

Li, J. et al., 2014. An integrated catalog of reference genes in the human gut microbiome. *Nature Biotechnology*, 32(8), pp.834–841.

Li, X. et al., 2013. Toll-like receptor 4 increases intestinal permeability through up-regulation of membrane PKC activity in alcoholic steatohepatitis. *Alcohol*, 47(6), pp.459–465.

Li, Y. et al., 2011. Exogenous Stimuli Maintain Intraepithelial Lymphocytes via Aryl Hydrocarbon Receptor Activation. *Cell*, 147(3), pp.629–640.

Liu, Z. et al., 2018. Aryl hydrocarbon receptor activation maintained the intestinal epithelial barrier function through Notch1 dependent signaling pathway. *International Journal of Molecular Medicine*, 41(3), pp.1560–1572.

Lowe, M.M. et al., 2014. Identification of cinnabarinic acid as a novel endogenous aryl hydrocarbon receptor ligand that drives IL-22 production M. Platten, ed. *PLoS ONE*, 9(2), p.e87877.

Lozupone, C.A. et al., 2012. Diversity, stability and resilience of the human gut microbiota. *Nature*, 489(7415), pp.220–230.

M'Koma, A.E., 2013. Inflammatory bowel disease: an expanding global health problem. *Clinical medicine insights. Gastroenterology*, 6, pp.33–47.

Ma, Q. & Baldwin, K.T., 2000. 2,3,7,8-tetrachlorodibenzo-p-dioxin-induced degradation of aryl hydrocarbon receptor (AhR) by the ubiquitin-proteasome pathway. Role of the transcription activaton and DNA binding of AhR. *The Journal of biological chemistry*, 275(12), pp.8432–8.

Ma, Q. & Whitlock, J.P., 1997. A novel cytoplasmic protein that interacts with the Ah receptor, contains tetratricopeptide repeat motifs, and augments the transcriptional response to 2,3,7,8-tetrachlorodibenzo-p-dioxin. *The Journal of biological chemistry*, 272(14), pp.8878–84.

Ma, Q. & Whitlock, J.P., 1996. The aromatic hydrocarbon receptor modulates the Hepa 1c1c7 cell cycle and differentiated state independently of dioxin. *Molecular and cellular biology*, 16(5), pp.2144–50.

MacPherson, A.J. & Uhr, T., 2004. Induction of Protective IgA by Intestinal Dendritic Cells Carrying Commensal Bacteria. *Science*, 303(5664), pp.1662–1665.

MacPherson, L. et al., 2014. Aryl hydrocarbon receptor repressor and TiPARP (ARTD14) use similar, but also distinct mechanisms to repress aryl hydrocarbon receptor signaling. *International journal of molecular sciences*, 15(5), pp.7939–57.

Magiatis, P. et al., 2013. Malassezia Yeasts Produce a Collection of Exceptionally Potent Activators of the Ah (Dioxin) Receptor Detected in Diseased Human Skin. *Journal of Investigative Dermatology*, 133(8), pp.2023–2030.

Manchester, D.K. et al., 1987. Ah receptor in human placenta: stabilization by molybdate and characterization of binding of 2,3,7,8-tetrachlorodibenzo-p-dioxin, 3-methylcholanthrene, and benzo(a)pyrene. *Cancer research*, 47(18), pp.4861–8.

Manichanh, C. et al., 2006. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut*, 55(2), pp.205–11.

Marcobal, A. et al., 2013. A refined palate: Bacterial consumption of host glycans in the gut. *Glycobiology*, 23(9), pp.1038–1046.

Martin-Gallausiaux, C. et al., Butyrate from commensal bacteria activates TGF-beta1 expression through

the transcription factor SP1 in human intestinal epithelial cells.

Maslowski, K.M. et al., 2009. Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. *Nature*, 461(7268), pp.1282–1286.

May, M.J. & Ghosh, S., 1998. Signal transduction through NF-kappa B. *Immunology today*, 19(2), pp.80–8.

Ménard, O. et al., 2008. Gnotobiotic mouse immune response induced by Bifidobacterium sp. strains isolated from infants. *Applied and environmental microbiology*, 74(3), pp.660–6.

Mexia, N. et al., 2015. Pityriazepin and other potent AhR ligands isolated from Malassezia furfur yeast. *Archives of Biochemistry and Biophysics*, 571, pp.16–20.

Meyer, B.K. et al., 1998. Hepatitis B virus X-associated protein 2 is a subunit of the unliganded aryl hydrocarbon receptor core complex and exhibits transcriptional enhancer activity. *Molecular and cellular biology*, 18(2), pp.978–88.

Meyer, B.K. & Perdew, G.H., 1999. Characterization of the AhR-hsp90-XAP2 core complex and the role of the immunophilin-related protein XAP2 in AhR stabilization. *Biochemistry*, 38(28), pp.8907–8917.

Miller, C.A., 1997. Expression of the human aryl hydrocarbon receptor complex in yeast. Activation of transcription by indole compounds. *The Journal of biological chemistry*, 272(52), pp.32824–9.

Mimura, J. et al., 1999. Identification of a novel mechanism of regulation of Ah (dioxin) receptor function. *Genes & development*, 13(1), pp.20–5.

Mimura, J. et al., 2003. Loss of teratogenic response to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in mice lacking the Ah (dioxin) receptor. *Genes to cells : devoted to molecular & cellular mechanisms*, 2(10), pp.645–654.

Mitchell, K.A. & Elferink, C.J., 2009. Timing is everything: Consequences of transient and sustained AhR activity. *Biochemical Pharmacology*, 77(6), pp.947–956.

Monteleone, I. et al., 2011. *Aryl hydrocarbon receptor-induced signals up-regulate IL-22 production and inhibit inflammation in the gastrointestinal tract: Commentary*,

Monticelli, L.A. et al., 2011. Innate lymphoid cells promote lung-tissue homeostasis after infection with influenza virus. *Nature immunology*, 12(11), pp.1045–54.

Moura-Alves, P. et al., 2014. AhR sensing of bacterial pigments regulates antibacterial defence. *Nature*, 512, pp.387–392.

Mueller, S. et al., 2006. Differences in Fecal Microbiota in Different European Study Populations in Relation to Age, Gender, and Country: a Cross-Sectional Study. *Applied and Environmental Microbiology*, 72(2), pp.1027–1033.

Murray, I.A. & Perdew, G.H., 2017. Ligand activation of the Ah receptor contributes to gastrointestinal homeostasis. *Current opinion in toxicology*, 2, pp.15–23.

Nebert, D.W. et al., 1969. The in vivo and in vitro induction of aryl hydrocarbon hydroxylase in mammalian cells of different species, tissues, strains, and developmental and hormonal states. *Archives of biochemistry and biophysics*, 134(1), pp.76–89.

Nebert, D.W., Wikvall, K. & Miller, W.L., 2013. Human cytochromes P450 in health and disease. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 368(1612), p.20120431.

Neill, D.R. et al., 2010. Nuocytes represent a new innate effector leukocyte that mediates type-2 immunity. *Nature*, 464(7293), pp.1367–1370.

Nelson, D.R. et al., 2004. Comparison of cytochrome P450 (CYP) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative-splice

variants. *Pharmacogenetics*, 14(14), pp.1–181.

Nepelska, M. et al., 2017. Commensal gut bacteria modulate phosphorylation-dependent PPARγ transcriptional activity in human intestinal epithelial cells. *Scientific Reports*, 7, p.43199.

Neutra, M.R., Frey, A. & Kraehenbuhl, J.P., 1996. Epithelial M cells: gateways for mucosal infection and immunization. *Cell*, 86(3), pp.345–8.

Nguyen, N.T. et al., 2010. Aryl hydrocarbon receptor negatively regulates dendritic cell immunogenicity via a kynurenine-dependent mechanism. *Proceedings of the National Academy of Sciences*, 107(46), pp.19961–19966.

Nibali, L. & Henderson, B., 2016. *The Human Microbiota and Chronic Disease. Dysbiosis as a cause of human pathology.*, Wiley-Blackwell.

Nigro, G. et al., 2014. The cytosolic bacterial peptidoglycan sensor Nod2 affords stem cell protection and links microbes to gut epithelial regeneration. *Cell host & microbe*, 15(6), pp.792–8.

Nishiumi, S., Yoshida, K.-I. ichi & Ashida, H., 2007. Curcumin suppresses the transformation of an aryl hydrocarbon receptor through its phosphorylation. *Archives of Biochemistry and Biophysics*, 466(2), pp.267–273.

Noverr, M.C. et al., 2005. Development of allergic airway disease in mice following antibiotic therapy and fungal microbiota increase: role of host genetics, antigen, and interleukin-13. *Infection and immunity*, 73(1), pp.30–8.

Ochoa-Reparaz, J. et al., 2010. Central Nervous System Demyelinating Disease Protection by the Human Commensal Bacteroides fragilis Depends on Polysaccharide A Expression. *The Journal of Immunology*, 185(7), pp.4101–4108.

Ohtake, F. et al., 2007. Dioxin receptor is a ligand-dependent E3 ubiquitin ligase. *Nature*, 446(7135), pp.562–566.

Ohtake, F., Fujii-Kuriyama, Y. & Kato, S., 2009. AhR acts as an E3 ubiquitin ligase to modulate steroid receptor functions. *Biochemical Pharmacology*, 77(4), pp.474–484.

Oshima, M. et al., 2009. SUMO modification regulates the transcriptional repressor function of aryl hydrocarbon receptor repressor. *The Journal of biological chemistry*, 284(17), pp.11017–26.

Otte, J.-M., Cario, E. & Podolsky, D.K., 2004. Mechanisms of cross hyporesponsiveness to Toll-like receptor bacterial ligands in intestinal epithelial cells. *Gastroenterology*, 126(4), pp.1054–70.

Owen, R.L., Piazza, A.J. & Ermak, T.H., 1991. Ultrastructural and cytoarchitectural features of lymphoreticular organs in the colon and rectum of adult BALB/c mice. *American Journal of Anatomy*, 190(1), pp.10–18.

Pennington, R.J., 1952. The metabolism of short-chain fatty acids in the sheep. I. Fatty acid utilization and ketone body production by rumen epithelium and other tissues. *The Biochemical journal*, 51(2), pp.251–8.

Perdew, G.H. & Bradfield, C.A., 1996. Mapping the 90 kDa heat shock protein binding region of the Ah receptor. *Biochemistry and molecular biology international*, 39(3), pp.589–93.

Peters, J.C., 1991. Tryptophan nutrition and metabolism: an overview. *Advances in experimental medicine and biology*, 294, pp.345–58.

Peters, J.M. & Wiley, L.M., 1995. Evidence That Murine Preimplantation Embryos Express Aryl Hydrocarbon Receptor. *Toxicology and Applied Pharmacology*, 134(2), pp.214–221.

Petkov, P.I. et al., 2010. Mechanism-based common reactivity pattern (COREPA) modelling of aryl hydrocarbon receptor binding affinity. *SAR and QSAR in environmental research*, 21(1), pp.187–214.

Petrulis, J.R. et al., 2003. The hsp90 Co-chaperone XAP2 Alters Importin beta Recognition of the Bipartite Nuclear Localization Signal of the Ah Receptor and Represses Transcriptional Activity. *Journal of Biological Chemistry*, 278(4), pp.2677–2685.

Poland, A. & Glover, E., 1990. Characterization and strain distribution pattern of the murine Ah receptor specified by the Ahd and Ahb-3 alleles. *Molecular pharmacology*, 38(3), pp.306–12.

Poland, A. & Glover, E., 1973. Chlorinated dibenzo-p-dioxins: potent inducers of delta-aminolevulinic acid synthetase and aryl hydrocarbon hydroxylase. II. A study of the structure-activity relationship. *Molecular pharmacology*, 9(6), pp.736–47.

Poland, A. & Knutson, J.C., 1982. 2,3,7,8-Tetrachlorodibenzo- *p* -Dioxin and Related Halogenated Aromatic Hydrocarbons: Examination of the Mechanism of Toxicity. *Annual Review of Pharmacology and Toxicology*, 22(1), pp.517–554.

Poland, A.P. et al., 1974. Genetic expression of aryl hydrocarbon hydroxylase activity. Induction of monooxygenase activities and cytochrome P1-450 formation by 2,3,7,8 tetrachlorodibenzo p dioxin in mice genetically "nonresponsive" to other aromatic hydrocarbons. *Journal of Biological Chemistry*, 249(17), pp.5599–5606.

Pollenz, R.S., Sattler, C.A. & Poland, A., 1994. The aryl hydrocarbon receptor and aryl hydrocarbon receptor nuclear translocator protein show distinct subcellular localizations in Hepa 1c1c7 cells by immunofluorescence microscopy. *Molecular pharmacology*, 45(3), pp.428–438.

Pongratz, I. et al., 1998. Role of the PAS domain in regulation of dimerization and DNA binding specificity of the dioxin receptor. *Molecular and cellular biology*, 18(7), pp.4079–88.

Pongratz, I., Mason, G.G. & Poellinger, L., 1992. Dual roles of the 90-kDa heat shock protein hsp90 in modulating functional activities of the dioxin receptor. Evidence that the dioxin receptor functionally belongs to a subclass of nuclear receptors which require hsp90 both for ligand binding activity and repression of intrinsic DNA binding activity. *The Journal of biological chemistry*, 267(19), pp.13728–34.

Porter, E.M. et al., 2002. The multifaceted Paneth cell. *Cellular and molecular life sciences : CMLS*, 59(1), pp.156–70.

Le Poul, E. et al., 2003. Functional characterization of human receptors for short chain fatty acids and their role in polymorphonuclear cell activation. *The Journal of biological chemistry*, 278(28), pp.25481–9.

Puga, A., Ma, C. & Marlowe, J.L., 2009. The aryl hydrocarbon receptor cross-talks with multiple signal transduction pathways. *Biochemical Pharmacology*, 77(4), pp.713–722.

Purushe, J. et al., 2010. Comparative Genome Analysis of Prevotella ruminicola and Prevotella bryantii: Insights into Their Environmental Niche. *Microbial Ecology*, 60(4), pp.721–729.

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

Qiu, J. et al., 2013. Group 3 Innate Lymphoid Cells Inhibit T-Cell-Mediated Intestinal Inflammation through Aryl Hydrocarbon Receptor Signaling and Regulation of Microflora. *Immunity*, 39(2), pp.386–399.

Qiu, J. et al., 2012. The Aryl Hydrocarbon Receptor Regulates Gut Immunity through Modulation of Innate Lymphoid Cells. *Immunity*, 36(1), pp.92–104.

Qui, J. et al., 2013. Aryl Hydrocarbon Receptor Promotes RORγt+ ILCs and Controls Intestinal Immunity and Inflammation. *Semin Immunopathol*, 35(6), pp.657–670.

Quintana, F.J. et al., 2008. Control of Treg and TH17 cell differentiation by the aryl hydrocarbon receptor. *Nature*, 453(7191), pp.65–71.

Rannug, A. et al., 1987. Certain photooxidized derivatives of tryptophan bind with very high affinity to the Ah receptor and are likely to be endogenous signal substances. *The Journal of biological chemistry*, 262(32), pp.15422–7.

Rannug, U. et al., 1995. Structure elucidation of two tryptophan-derived, high affinity Ah receptor ligands. *Chemistry & biology*, 2(12), pp.841–5.

Rasmussen, M.K. et al., 2016. Skatole (3-Methylindole) Is a Partial Aryl Hydrocarbon Receptor Agonist and Induces CYP1A1/2 and CYP1B1 Expression in Primary Human Hepatocytes. *PloS one*, 11(5), p.e0154629.

Rehman, A. et al., 2011. Nod2 is essential for temporal development of intestinal microbial communities. *Gut*, 60(10), pp.1354–1362.

Rescigno, M. et al., 2001. Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria. *Nature immunology*, 2(4), pp.361–7.

Ridaura, V.K. et al., 2013. Gut Microbiota from Twins Discordant for Obesity Modulate Metabolism in Mice. *Science*, 341(6150), pp.1241214–1241214.

Romano, K.A. et al., 2015. Intestinal Microbiota Composition Modulates Choline Bioavailability from Diet and Accumulation of the Proatherogenic Metabolite Trimethylamine-N-Oxide. *mBio*, 6(2).

Romeo, E. et al., 2006. Luminal kidney and intestine SLC6 amino acid transporters of B0AT-cluster and their tissue distribution in Mus musculus. *American journal of physiology. Renal physiology*, 290(2), pp.F376–F383.

Round, J.L. et al., 2011. The Toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota. *Science (New York, N.Y.)*, 332(6032), pp.974–7.

Russell, S.L. et al., 2012. Early life antibiotic-driven changes in microbiota enhance susceptibility to allergic asthma. *EMBO reports*, 13(5), pp.440–447.

Safe, S., 1990. Polychlorinated Biphenyls (PCBs), Dibenzo-p-Dioxins (PCDDs), Dibenzofurans (PCDFs), and Related Compounds: Environmental and Mechanistic Considerations Which Support the Development of Toxic Equivalency Factors (TEFs). *Critical Reviews in Toxicology*, 21(1), pp.51–88.

Sakurai, S., Shimizu, T. & Ohto, U., 2017. The crystal structure of the AhRR–ARNT heterodimer reveals the structural basis of the repression of AhR-mediated transcription. *Journal of Biological Chemistry*, 292(43), pp.17609–17616.

Savinova, O. V, Hoffmann, A. & Ghosh, G., 2009. The Nfκb1 and Nfκb2 proteins p105 and p100 function as the core of high-molecular-weight heterogeneous complexes. *Molecular cell*, 34(5), pp.591–602.

Schaldach, C.M., Riby, J. & Bjeldanes, L.F., 1999. Lipoxin A $_4$: A New Class of Ligand for the Ah Receptor. *Biochemistry*, 38(23), pp.7594–7600.

Schiering, C. et al., 2017. Feedback control of AHR signalling regulates intestinal immunity. *Nature*, 542(7640), pp.242–245.

Schilderink, R., Verseijden, C. & de Jonge, W.J., 2013. Dietary Inhibitors of Histone Deacetylases in Intestinal Immunity and Homeostasis. *Frontiers in Immunology*, 4, p.226.

Schroeder, J.C. et al., 2010. The Uremic Toxin 3-Indoxyl Sulfate Is a Potent Endogenous Agonist for the Human Aryl Hydrocarbon Receptor. *Biochemistry*, 49(2), pp.393–400.

Segura, E. et al., 2013. Human Inflammatory Dendritic Cells Induce Th17 Cell Differentiation. *Immunity*, 38(2), pp.336–348.

Seliskar, M. & Rozman, D., 2007. Mammalian cytochromes P450—Importance of tissue specificity. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 1770(3), pp.458–466.

Seok, S.-H. et al., 2017. Structural hierarchy controlling dimerization and target DNA recognition in the AHR transcriptional complex. *PNAS*, 114(21), pp.5431–5436.

Shanahan, F., 2002. The host-microbe interface within the gut. *Best practice & research. Clinical gastroenterology*, 16(6), pp.915–31.

Shen, E.S. & Whitlock, J.P., 1992. Protein-DNA interactions at a dioxin-responsive enhancer. Mutational analysis of the DNA-binding site for the liganded Ah receptor. *Journal of Biological Chemistry*, 267(10), pp.6815–6819.

Shimba, S. et al., 2003. Transcriptional regulation of the AhR gene during adipose differentiation. *Biological & pharmaceutical bulletin*, 26(9), pp.1266–71.

Shin, J.H. et al., 2013. Modulation of natural killer cell antitumor activity by the aryl hydrocarbon receptor. *Proceedings of the National Academy of Sciences*, 110(30), pp.12391–12396.

Shulzhenko, N. et al., 2011. Crosstalk between B lymphocytes, microbiota and the intestinal epithelium governs immunity versus metabolism in the gut. *Nature Medicine*, 17(12), pp.1585–1593.

Simon, G.M., Cheng, J. & Gordon, J.I., 2012. Quantitative assessment of the impact of the gut microbiota on lysine -acetylation of host proteins using gnotobiotic mice. *Proceedings of the National Academy of Sciences*, 109(28), pp.11133–11138.

Sinal, C.J. & Bend, J.R., 1997. Aryl hydrocarbon receptor-dependent induction of cyp1a1 by bilirubin in mouse hepatoma hepa 1c1c7 cells. *Molecular pharmacology*, 52(4), pp.590–9.

Van der Sluis, M. et al., 2006. Muc2-Deficient Mice Spontaneously Develop Colitis, Indicating That MUC2 Is Critical for Colonic Protection. *Gastroenterology*, 131(1), pp.117–129.

Smith, B.W. et al., 2013. The aryl hydrocarbon receptor directs hematopoietic progenitor cell expansion and differentiation. *Blood*, 122(3), pp.376–85.

Smith, E.A. & Macfarlane, G.T., 1997. Formation of Phenolic and Indolic Compounds by Anaerobic Bacteria in the Human Large Intestine. *Microbial Ecology*, 33(3), pp.180–188.

Smith, P.M. et al., 2013. The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science (New York, N.Y.)*, 341(6145), pp.569–73.

Sokol, H. et al., 2008. Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. *Proceedings of the National Academy of Sciences of the United States of America*, 105(43), pp.16731–6.

Song, J. et al., 2002. A ligand for the aryl hydrocarbon receptor isolated from lung. *Proceedings of the National Academy of Sciences of the United States of America*, 99(23), pp.14694–9.

Sonnenberg, G.F. et al., 2012. Innate Lymphoid Cells Promote Anatomical Containment of Lymphoid-Resident Commensal Bacteria. *Science*, 336(6086), pp.1321–1325.

Sonowal, R. et al., 2017. Indoles from commensal bacteria extend healthspan. *Proceedings of the National Academy of Sciences*, 114(36), pp.E7506–E7515.

Soshilov, A.A. & Denison, M.S., 2014. Ligand promiscuity of aryl hydrocarbon receptor agonists and antagonists revealed by site-directed mutagenesis. *Molecular and cellular biology*, 34(9), pp.1707–1719.

Spencer, J. et al., 1991. Gamma/delta T cells and the diagnosis of coeliac disease. *Clinical and experimental immunology*, 85(1), pp.109–13.

Stockinger, B. et al., 2014. The Aryl Hydrocarbon Receptor: Multitasking in the Immune System. *Annual Review of Immunology*, 32(1), pp.403–432.

Swanson, H.I., 2002. DNA binding and protein interactions of the AHR/ARNT heterodimer that facilitate

gene activation. *Chemico-biological interactions*, 141(1–2), pp.63–76.

Swanson, H.I. & Yang, J.H., 1998. The aryl hydrocarbon receptor interacts with transcription factor IIB. *Molecular pharmacology*, 54(4), pp.671–7.

Tait Wojno, E.D. & Artis, D., 2012. Innate Lymphoid Cells: Balancing Immunity, Inflammation, and Tissue Repair in the Intestine. *Cell Host & Microbe*, 12(4), pp.445–457.

Takamura, T. et al., 2011. Lactobacillus bulgaricus OLL1181 activates the aryl hydrocarbon receptor pathway and inhibits colitis. *Immunology and Cell Biology*, 89(7), pp.817–822.

Tan, M. et al., 2011. Identification of 67 Histone Marks and Histone Lysine Crotonylation as a New Type of Histone Modification. *Cell*, 146(6), pp.1016–1028.

Tan, X. et al., 2017. Indoxyl sulfate, a valuable biomarker in chronic kidney disease and dialysis. *Hemodialysis International*, 21(2), pp.161–167.

Tap, J. et al., 2009. Towards the human intestinal microbiota phylogenetic core. *Environmental microbiology*, 11(10), pp.2574–84.

Thangaraju, M. et al., 2009. Transport by SLC5A8 with subsequent inhibition of histone deacetylase 1 (HDAC1) and HDAC3 underlies the antitumor activity of 3-bromopyruvate. *Cancer*, 115(20), pp.4655–4666.

Thia, K.T. et al., 2008. An Update on the Epidemiology of Inflammatory Bowel Disease in Asia. *The American Journal of Gastroenterology*, 103(12), pp.3167–3182.

Thomas, L. V. & Ockhuizen, T., 2012. New insights into the impact of the intestinal microbiota on health and disease: a symposium report. *British Journal of Nutrition*, 107(S1), pp.S1–S13.

Tian, Y. et al., 1999. Ah receptor and NF-kappaB interactions, a potential mechanism for dioxin toxicity. *The Journal of biological chemistry*, 274(1), pp.510–5.

Tijet, N. et al., 2006. Aryl hydrocarbon receptor regulates distinct dioxin-dependent and dioxin-independent gene batteries. *Molecular pharmacology*, 69(1), pp.140–53.

Tkachenko, A. et al., 2016. The Q-rich/PST domain of the AHR regulates both ligand-induced nuclear transport and nucleocytoplasmic shuttling. *Scientific Reports*, 6(1), p.32009.

Tsuchiya, Y. et al., 2003. Expression of aryl hydrocarbon receptor repressor in normal human tissues and inducibility by polycyclic aromatic hydrocarbons in human tumor-derived cell lines. *Toxicological Sciences*, 72(2), pp.253–259.

Turnbaugh, P.J. et al., 2009. A core gut microbiome in obese and lean twins. *Nature*, 457(7228), pp.480–4.

Veldhoen, M. et al., 2009. Natural agonists for aryl hydrocarbon receptor in culture medium are essential for optimal differentiation of Th17 T cells. *The Journal of experimental medicine*, 206(1), pp.43–49.

Veldhoen, M. et al., 2008. The aryl hydrocarbon receptor links TH17-cell-mediated autoimmunity to environmental toxins. *Nature*, 453(7191), pp.106–109.

Venkatesh, M. et al., 2014. Symbiotic Bacterial Metabolites Regulate Gastrointestinal Barrier Function via the Xenobiotic Sensor PXR and Toll-like Receptor 4. *Immunity*, 41(2), pp.296–310.

Vikström Bergander, L. et al., 2012. Tryptamine serves as a proligand of the AhR transcriptional pathway whose activation is dependent of monoamine oxidases. *Molecular endocrinology (Baltimore, Md.)*, 26(9), pp.1542–51.

Villa, M. et al., 2017. Aryl hydrocarbon receptor is required for optimal B-cell proliferation. *The EMBO journal*, 36(1), pp.116–128.

Vogel, C.F. a et al., 2007. RelB, a new partner of aryl hydrocarbon receptor-mediated transcription. *Molecular endocrinology (Baltimore, Md.)*, 21(12), pp.2941–2955.

Vrzal, R. et al., 2017. Activated thyroid hormone receptor modulates dioxin-inducible aryl hydrocarbon receptor-mediated CYP1A1 induction in human hepatocytes but not in human hepatocarcinoma HepG2 cells. *Toxicology Letters*, 275, pp.77–82.

Wang, F. et al., 2015. Gut Microbiota Community and Its Assembly Associated with Age and Diet in Chinese Centenarians. *Journal of Microbiology and Biotechnology*, 25(8), pp.1195–1204.

Wang, I.-K. et al., 2015. The effect of probiotics on serum levels of cytokine and endotoxin in peritoneal dialysis patients: a randomised, double-blind, placebo-controlled trial. *Beneficial Microbes*, 6(4), pp.423–430.

Wang, Z. et al., 2011. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature*, 472, pp.57–63.

Wansa, K.D.S.A. et al., 2003. The AF-1 domain of the orphan nuclear receptor NOR-1 mediates trans-activation, coactivator recruitment, and activation by the purine anti-metabolite 6-mercaptopurine. *The Journal of biological chemistry*, 278(27), pp.24776–90.

Wei, Y.D. et al., 1998. Rapid and transient induction of CYP1A1 gene expression in human cells by the tryptophan photoproduct 6-formylindolo[3,2- b]carbazole. *Chemico-Biological Interactions*, 110(1–2), pp.39–55.

Wellen, K.E. & Hotamisligil, G.S., 2005. Inflammation, stress, and diabetes. *The Journal of clinical investigation*, 115(5), pp.1111–9.

Wernet, M.F. et al., 2006. Stochastic spineless expression creates the retinal mosaic for colour vision. *Nature*, 440(7081), pp.174–180.

Wikoff, W.R. et al., 2009. Metabolomics analysis reveals large effects of gut microflora on mammalian blood metabolites. *Proceedings of the National Academy of Sciences of the United States of America*, 106(10), pp.3698–3703.

Wong, J.M.W. et al., 2006. Colonic health: fermentation and short chain fatty acids. *Journal of clinical gastroenterology*, 40(3), pp.235–43.

Wong, P.S. et al., 2010. Arylhydrocarbon receptor activation in NCI-H441 cells and C57BL/6 mice: possible mechanisms for lung dysfunction. *American journal of respiratory cell and molecular biology*, 42(2), pp.210–7.

Wu, G.D. et al., 2011. Linking Long-Term Dietary Patterns with Gut Microbial Enterotypes. *Science*, 334(6052), pp.105–108.

Wu, H.-J. et al., 2010. Gut-Residing Segmented Filamentous Bacteria Drive Autoimmune Arthritis via T Helper 17 Cells. *Immunity*, 32(6), pp.815–827.

Xie, Z. et al., 2016. Metabolic Regulation of Gene Expression by Histone Lysine β-Hydroxybutyrylation. *Molecular Cell*, 62(2), pp.194–206.

Yamaori, S. et al., 2015. Cannabidiol induces expression of human cytochrome P450 1A1 that is possibly mediated through aryl hydrocarbon receptor signaling in HepG2 cells. *Life Sciences*, 136, pp.87–93.

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

Ye, J. et al., 2017. The Aryl Hydrocarbon Receptor Preferentially Marks and Promotes Gut Regulatory T Cells. *Cell reports*, 21(8), pp.2277–2290.

Young, V.B. & Schmidt, T.M., 2004. Antibiotic-associated diarrhea accompanied by large-scale alterations

in the composition of the fecal microbiota. *Journal of clinical microbiology*, 42(3), pp.1203–6.

Yu, M. et al., 2018. Aryl Hydrocarbon Receptor Activation Modulates Intestinal Epithelial Barrier Function by Maintaining Tight Junction Integrity. *International Journal of Biological Sciences*, 14(1), pp.69–77.

Zeisel, S.H. & Da Costa, K.A., 2009. Choline: An essential nutrient for public health. *Nutrition Reviews*, 67, pp.615–623.

Zelante, T. et al., 2013. Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22. *Immunity*, 39(6), pp.372–385.

Zenewicz, L.A. et al., 2008. Innate and adaptive interleukin-22 protects mice from inflammatory bowel disease. *Immunity*, 29(6), pp.947–57.

Zhang, S., Qin, C. & Safe, S.H., 2003. Flavonoids as aryl hydrocarbon receptor agonists/antagonists: effects of structure and cell context. *Environmental health perspectives*, 111(16), pp.1877–82.

Zheng, L. et al., 2017. Microbial-Derived Butyrate Promotes Epithelial Barrier Function through IL-10 Receptor–Dependent Repression of Claudin-2. *The Journal of Immunology*, p.ji1700105.

Zimmer, J. et al., 2012. A vegan or vegetarian diet substantially alters the human colonic faecal microbiota. *European Journal of Clinical Nutrition*, 66(1), pp.53–60.

Zindl, C.L. et al., 2013. IL-22-producing neutrophils contribute to antimicrobial defense and restitution of colonic epithelial integrity during colitis. *Proceedings of the National Academy of Sciences of the United States of America*, 110(31), pp.12768–73.

Zoetendal, E.G. et al., 2002. Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. *Applied and environmental microbiology*, 68(7), pp.3401–7.

Zoetendal, E.G. et al., 2012. The human small intestinal microbiota is driven by rapid uptake and conversion of simple carbohydrates. *The ISME Journal*, 6(7), pp.1415–1426.

Scientific Summary

The intestinal microbiota shapes the host physiology through the production of various metabolites. The transcription factor Aryl hydrocarbon Receptor (AhR) emerges as an actor of the host-microbiota crosstalk. Indeed, numerous bacterial molecules are described to activate AhR pathway and being involved in the intestinal homeostasis. Among them, indoles and other tryptophan-derived metabolites produced by commensal and probiotic strains were described for protecting mice from induced colitis and their abundance being inversely correlated with inflammatory bowel diseases (IBD) in humans. However, the current knowledge on bacterial molecules activating AhR pathway, is still limited, considering the complexity of the intestinal microbial community. By screening a collection of commensal bacteria on human intestinal epithelial cell lines, we identified microbial modulators of AhR pathway. The use of cells expressing an AhR reporter system allowed the identification of activating bacterial strains and discriminate different mechanisms of action. Firstly, bacteria producing short chain fatty acids (SCFA) emerged as strong activators of AhR pathway and we showed for the first time that butyrate acts as a novel AhR ligand. Additionally, some bacteria not predicted to produce butyrate nor indoles, were identified as AhR activators in our screen. Among them, some species belonging to the Actinobacteria phylum seem a promising group of AhR activators, through the production of a microbial metabolite not yet identified. In conclusion, this work sheds light on a novel role of butyrate as AhR ligand and introduces a newly potential family of AhR activator produced by Actinobacteria.

Key words: AhR; microbial metabolites; butyrate; Actinobacteria; intestinal microbiota; intestinal epithelial cells.

Résumé de thèse

Le microbiote intestinal joue un rôle fondamental dans la modulation du métabolisme et du système immunitaire de l'hôte à travers la production de métabolites. Le récepteur aux hydrocarbures aromatiques (AhR) est un acteur important dans l'interaction entre le microbiote et l'hôte. En effet, plusieurs métabolites microbiens ont été décrits comme activateurs de la voie AhR et impliqués dans l'homéostasie intestinale. Parmi eux, les indoles et autres métabolites dérivés du tryptophane, produits par des bactéries commensales ou probiotiques, ont été décrites pour protéger les souris lors de colites induites alors que leur présence est peu détectée chez les patients atteints de maladies inflammatoires chroniques de l'intestin (MICI). Le criblage d'une collection de bactéries commensales sur des lignées de cellules épithéliales intestinales a permis d'identifier des souches bactériennes modulatrices de la voie AhR. L'utilisation des cellules exprimant un système rapporteur AhR, a révélé des activateurs bactériens et mis en évidence différents mécanismes d'action. Nous avons montré que des bactéries productrices d'acides gras à chaine courte (AGCC) sont de forts activateurs de la voie AhR et que le butyrate semblait être un nouveau ligand d'AhR. De plus, nous avons identifié des Actinobacteries, non décrites pour produire du butyrate ou des indoles, comme activateurs de la voie AhR via la production d'un métabolite microbien non identifié à ce jour. En conclusion, ce travail illustre un nouveau rôle fonctionnel du butyrate comme ligand d'AhR et montre l'existence d'une nouvelle famille de métabolites microbiens produits par des Actinobacteries et activatrice de la voie AhR.

Mots clés: AhR ; métabolites microbiens ; butyrate ; Actinobacteria ; microbiote intestinale ; cellules épithéliales intestinales.