

(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 2014202021 B2**

(54) Title
Polypeptide and immune modulation

(51) International Patent Classification(s)
A61K 35/74 (2006.01) **A61P 37/02** (2006.01)
A61K 31/713 (2006.01)

(21) Application No: **2014202021** (22) Date of Filing: **2014.04.10**

(30) Priority Data

(31)	Number	(32)	Date	(33)	Country
	1306536.2		2013.04.10		GB

(43) Publication Date: **2014.10.30**

(43) Publication Journal Date: **2014.10.30**

(44) Accepted Journal Date: **2019.02.28**

(71) Applicant(s)
4D Pharma Research Limited

(72) Inventor(s)
Kelly, Denise;Patterson, Angela;Monnais, Edouard;Mulder, Imke

(74) Agent / Attorney
FB Rice Pty Ltd, L 23 44 Market St, Sydney, NSW, 2000, AU

(56) Related Art
DUNCAN SH, et al., 'Roseburia intestinalis sp. nov., a novel saccharolytic, butyrate-producing bacterium from human faeces,' International Journal of Systemic and Evolutionary Microbiology, (2002) vol 52, pp 1615-1620.
NEVILLE AB. 'Functional genomics of motile commensal intestinal bacteria,' PhD Thesis, University College Cork, (15 January 2013), Ann Neville, Cork, Republic of Ireland, XP055133403.
WO 2013/050792 A1

ABSTRACT

Polypeptide and Immune Modulation

5 The present invention relates to *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell, including bacteria, comprising said vector, and/or a host cell, including bacteria, comprising said polynucleotide sequence, for use in modulating the inflammation of a tissue or an organ in a subject.

POLYPEPTIDE and IMMUNE MODULATION

FIELD OF THE INVENTION

5 The present invention relates to *Roseburia* flagellins, and/or polynucleotide sequences encoding said *Roseburia* flagellins, and/or vectors comprising said polynucleotide sequences, and/or host cells, including bacteria, comprising said vectors, and/or host cells, including bacteria, comprising said polynucleotide sequences, for various therapeutic and nutritional uses.

BACKGROUND

10
15 The human intestine is thought to be sterile *in utero*, but is exposed to a large variety of maternal and environmental microbes immediately after birth. Thereafter, a dynamic period of microbial colonization and succession occurs which is influenced by factors such as delivery mode, environment, diet and host genotype, all of which impact upon the composition of the gut microbiota, particularly during early life. Subsequently, the microbiota stabilizes and becomes adult-like (Spor, Koren & Ley 2011). The human gut microbiota contains more than 500-1000 different phylotypes belonging essentially to two major
20 bacterial divisions, the Bacteroidetes and the Firmicutes (Eckburg *et al.* 2005). The successful symbiotic relationships arising from bacterial colonization of the human gut have yielded a wide variety of metabolic, structural, protective and other beneficial functions. The enhanced metabolic activities of the colonized gut ensure that otherwise indigestible dietary components are degraded with release of byproducts providing an important nutrient source
25 for the host. Similarly, the immunological importance of the gut microbiota is well-recognized and is exemplified in germfree animals which have an impaired immune system that is functionally reconstituted following the introduction of commensal bacteria (Macpherson *et al.* 2001, Macpherson, Martinic & Harris 2002, Mazmanian *et al.* 2005).

30 In sharp contrast to the production of secretory intestinal IgA, which is influenced by microbial colonization *per se* (Chung, Kasper 2010, Macpherson 2006), T cell development and differentiation seem to require colonization by specific commensal micro-organisms. Clostridium species, and in particular the spore-forming segmented filamentous bacteria (SFB), appear to be a potent stimulus for the differentiation and maturation of intestinal and
35 colonic Th1, Th17 and Tregs (Gaboriau-Routhiau *et al.* 2009, Ivanov *et al.* 2009). Recent studies have now demonstrated that other gut bacteria, including those of Clostridium clusters IV and XIVa and the Altered Schaedler Flora (ASF), can induce *de novo* generation

of Tregs while mono-colonization with *Bacteroides fragilis* can correct the Th1/Th2 imbalance in germfree mice by promoting the expansion of Tregs (Mazmanian *et al.* 2005, Geuking *et al.* 2011, Atarashi *et al.* 2011). These data infer important immune-regulatory effects of other resident gut bacteria. Clearly, the effects of commensal bacteria on T cell differentiation pathways is variable and as postulated previously may be influenced by the array of TLR ligands found associated with specific bacteria (Nutsch, Hsieh 2012). For example, the mechanism by which SFB influences T cell responses is currently unknown, but recent genome studies confirming the presence of flagellin genes suggest that innate responses mediated through TLR5-flagellin interactions may be important (Prakash *et al.* 2011, Szczesnak *et al.* 2011). Furthermore, the Treg-enhancing effects of *B. fragilis* have been linked with PSA and mediation by TLR2 signalling events (Round *et al.* 2011).

Dramatic changes in microbiota composition have been documented in gastrointestinal disorders such as inflammatory bowel disease (IBD). For example, the levels of Clostridium cluster XIVa bacteria are reduced in IBD patients whilst numbers of *E. coli* are increased, suggesting a shift in the balance of symbionts and pathobionts within the gut (Frank *et al.* 2007, Scanlan *et al.* 2006, Kang *et al.* 2010, Machiels K. *et al.* 2013). Interestingly, this microbial dysbiosis is also associated with imbalances in T effector cell populations.

Roseburia belongs to the phylogenetic cluster XIVa of the Firmicutes phylum. Currently, within the *Roseburia* genus, five species have been identified and characterised: *Roseburia cecicola*, *Roseburia faecis*, *Roseburia hominis*, *Roseburia intestinalis*, *Roseburia inulinivorans* (Stanton and Savage 1983, Duncan *et al.* 2006). The bacteria are flagellated commensal anaerobes and are also major butyrate producers (Duncan *et al.* 2006). Although precise numbers of these bacteria colonizing the human gut have not been accurately estimated, *Roseburia* spp. are dominant in the healthy human gut and are under-represented in IBD patients (Machiels K. *et al.* 2013).

The roles of bacterial genes, in particular flagellin, participating in colonization and adaptation to the murine gut, as well as the host immune genes responding to colonization by *Roseburia* bacteria are disclosed. The inventors show that specific flagellin structures of *Roseburia* such as *R. hominis* and *R. intestinalis* induce distinct signalling responses in both epithelial cells and dendritic cells relative to other flagellate enteric bacteria. The importance of the TLR5-*Roseburia*, such as TLR5-*R. hominis*, interactions in directing the host adaptive response, in particular Treg responses is demonstrated. The complete genome sequence and annotation for *R. hominis* described herein is shown in GenBank under accession number CP003040 (version 1). For *R. intestinalis* (GenBank Accession Number for 16S

rRNA gene strain L1-82: AJ312385) described herein the reference genome sequence is shown in GenBank under accession number ABYJ02000000 (version 2) and consists of sequences ABYJ02000001- ABYJ02000409.

5 STATEMENTS OF INVENTION

Surprisingly, the present inventors found that *Roseburia* flagellin proteins are important in modulating the immune response.

10 In addition, the present inventors surprisingly found that the flagellin proteins derived or derivable from *Roseburia hominis* or *Roseburia intestinalis* are important in modulating the immune response.

15 The present invention relates to a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in modulating the inflammation of a tissue or an organ (such as the intestine) in a subject.

20 In another aspect, the present invention relates to a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species
25 *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in modulating the production of T cells (e.g. regulatory T cells such as regulatory T cells capable of expressing TLR5) in a subject.

30 The present invention relates to a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in restoring immunological tolerance.

35 In a further aspect, the present invention relates to a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell

comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in regulating the immune system of a subject.

5 In another aspect, the present invention relates to a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in treating a
10 disorder in a subject, wherein said disorder is an inflammatory disorder and/or an autoimmune disorder.

The present invention, in another aspect, relates to a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said
15 polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in modulating dendritic cells (such as bone marrow dendritic cells) and/or epithelial cells in a tissue or an organ of a subject.

20 The present invention, in a further aspect, relates to a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species
25 *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in regulating the production of IL-10 and/ or TGF β in a cell or cells of a subject.

In a further aspect, the present invention relates to a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said
30 polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in regulating the production of cell surface markers involved in immune responses and antigen recognition such as CD40, I-A/I-E, CD317/BST-2, CD103, CD80, CD86, CD83 and/or Siglec-H and/or
35 the species equivalent in a cell or cells of a subject.

In another aspect, the present invention relates to a *Roseburia* flagellin, and/or a

polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in regulating (e.g. downregulating) the expression of one of more Type I IFN genes (such as but not limited to one or more genes selected from the group consisting of IFN- β 1, IFN- β 3, *Ifi202b*, *Ifi203*, *IFI44*, *IFTI*, *MXI*, *OASI*, *OAS2*, *OAS3*, *OASL*, *Irf3* and *Irf4*) in a cell or cells of a subject.

The present invention, in a further aspect, relates to a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in regulating (e.g. downregulating) the expression of one or more pro-inflammatory genes (such as one or more genes selected from the group consisting of but not limited to IL1- β , IL4, IL5, IL6, IL8, IL12, IL13, IL17, IL21, IL22, IL23, IL27, IFN γ , CCL2, CCL3, CCL5, CCL20, CXCL5, CXCL10, CXCL12, CXCL13, and TNF- α) in a cell or cells of a subject.

The present invention, in another aspect, relates to a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in improving intestinal microbiota in a subject.

In another aspect, the present invention relates to a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in regulating appetite in a subject.

In a further aspect, the present invention relates to a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in regulating (e.g.

downregulating) the expression of the gene encoding cholecystokinin (Cck) and/or the expression of the gene encoding glucagon (Gcg) in a cell or cells of a subject.

The present invention, in a further aspect, relates to a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in improving alimentary canal health in a subject.

The present invention, in another aspect, relates to a pharmaceutical composition comprising a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and a pharmaceutically acceptable excipient, carrier or diluent.

In another aspect, the present invention relates to a nutritional supplement comprising a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and a nutritional acceptable excipient, carrier or diluent.

In a further aspect, the present invention relates to a feedstuff, food product, dietary supplement, or food additive comprising a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*).

The present invention, in a further aspect, relates to a process for producing a pharmaceutical composition according to the present invention, said process comprising admixing a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species

Roseburia intestinalis), with a pharmaceutically acceptable excipient, carrier or diluent; optionally said *Roseburia* flagellin, and/or said polynucleotide sequence, and/or said vector, and/or said host cell comprising said vector, and/or said host cell comprising said polynucleotide sequence, and/or said *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is encapsulated.

In a further aspect, the present invention relates to a process for producing a nutritional supplement according to the present invention, said process comprising admixing a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), with a nutritionally acceptable excipient, carrier or diluent; optionally said *Roseburia* flagellin, and/or said polynucleotide sequence, and/or said vector, and/or said host cell comprising said vector, and/or said host cell comprising said polynucleotide sequence, and/or said *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is encapsulated.

In another aspect, the present invention relates to a method of modulating the inflammation of a tissue or an organ (such as the intestine) in a subject, said method comprising administering to the subject a *Roseburia* flagellin, or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and wherein the inflammation of the tissue or organ (such as the intestine) in the subject is modulated.

The present invention, in another aspect, relates to a method of modulating the production of T cells (e.g. regulatory T cells such as regulatory T cells capable of expressing TLR5) in a subject, said method comprising administering a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and wherein the production of T cells (e.g. regulatory T cells such as regulatory T cells capable of expressing TLR5) in the subject is modulated.

The present invention, in a further aspect, relates to a method of regulating the immune system of a subject, said method comprising administering a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and wherein the immune system of the subject is regulated.

In a further aspect, the present invention relates to a method of treating a disorder in a subject, said method comprising administering a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), wherein said disorder is an inflammatory disorder and/or an autoimmune disorder.

In another aspect, the present invention relates to a method of modulating dendritic cells (such as bone marrow dendritic cells) and/or epithelial cells in a subject, said method comprising administering a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and wherein dendritic cells (such as bone marrow dendritic cells) and/or epithelial cells in the subject are modulated.

In a further aspect, the present invention relates to a method of regulating the production of IL-10 and/or TGF β in a cell or cells of a subject, said method comprising administering a *Roseburia* flagellin and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), to the subject, and wherein the production of IL-10 and/or TGF β in a cell or cells of the subject is regulated.

The present invention, in another aspect, relates to a method of regulating the production of cell surface markers involved in immune responses and antigen recognition such as CD40, I-A/I-E, CD317/BST-2, CD103, CD80, CD86, CD83 and/or Siglec-H and/or the species

equivalent in a cell or cells of a subject, said method comprising administering a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), to the subject, and wherein the production of cell surface markers involved in immune responses and antigen recognition such as CD40, I-A/I-E, CD317/BST-2, CD103, CD80, CD86, CD83 and/or Siglec-H and/or the species equivalent in a cell or cells of the subject is regulated.

In another aspect, the present invention relates to a method of regulating (e.g. downregulating) the expression of one of more Type I IFN genes (such as one or more genes selected from the group consisting of but not limited to IFN- β 1, IFN- β 3, *Ifi202b*, *Ifi203*, *IFI44*, *IFTI*, *MXI*, *OASI*, *OAS2*, *OAS3*, *OASL*, *Irf3* and *Irf4*) in a cell or cells of a subject, said method comprising administering a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), wherein the expression of one of more Type I IFN genes (such as one or more genes selected from the group consisting of but not limited to IFN- β 1, IFN- β 3, *Ifi202b*, *Ifi203*, *IFI44*, *IFTI*, *MXI*, *OASI*, *OAS2*, *OAS3*, *OASL*, *Irf3* and *Irf4*) in a cell or cells of the subject is regulated.

In a further aspect, the present invention relates to a method of regulating (e.g. downregulating) the expression of one or more pro-inflammatory genes (such as one or more genes selected from the group consisting of but not limited to IL1- β , IL4, IL5, IL6, IL8, IL12, IL13, IL17, IL21, IL22, IL23, IL27, IFN γ , CCL2, CCL3, CCL5, CCL20, CXCL5, CXCL10, CXCL12, CXCL13, and TNF- α) in a cell or cells of a subject, said method comprising administering a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), wherein the expression of one or more pro-inflammatory genes (such as one or more genes selected from the group consisting of but not limited to IL1- β , IL4, IL5, IL6, IL8, IL12, IL13, IL17, IL21, IL22, IL23, IL27, IFN γ , CCL3, CCL5, CCL20, CXCL5, CXCL10, CXCL12, CXCL13, and TNF- α) in a cell or cells of the subject is regulated.

The present invention, in a further aspect, relates to a method of improving intestinal

microbiota in a subject, said method comprising administering a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and wherein the intestinal microbiota in a subject is improved.

In another aspect, the present invention relates to a method of regulating appetite in a subject, said method comprising administering a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and wherein the appetite in the subject is regulated.

The present invention, in another aspect relates to a method of regulating (e.g. downregulating) the expression of the gene encoding cholecystokinin (Cck) and/or the expression of the gene encoding glucagon (Gcg) in a cell or cells of a subject, said method comprising administering a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and wherein the expression of the gene encoding cholecystokinin (Cck) and/or the expression of the gene encoding glucagon (Gcg) in a cell or cells of the subject is regulated.

In a further aspect, the present invention relates to a method of improving alimentary canal health in a subject, said method comprising administering a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and wherein alimentary canal health in a subject is improved.

The present invention, in a further aspect, relates to use of a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell

comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for the manufacture of a medicament for modulating the inflammation of a tissue or an organ (such as the intestine) in a subject.

The present invention, in another aspect, relates to use of a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for the manufacture of a medicament for modulating the production of T cells (e.g. regulatory T cells such as regulatory T cells capable of expressing TLR5) in a subject.

In a further aspect, the present invention relates to use of a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for the manufacture of a medicament for regulating the immune system of a subject.

The present invention, in a further aspect, relates to use of a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for the manufacture of a medicament for the treatment of a disorder in a subject, wherein said disorder is an inflammatory disorder and/or an autoimmune disorder.

The present invention, in another aspect, relates to use of a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for the manufacture of a medicament for modulating dendritic cells (such as bone marrow dendritic cells) and/or epithelial cells in a tissue or an organ of a subject.

In another aspect, the present invention relates to use of a *Roseburia* flagellin, and/or a

polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for the manufacture of a medicament for regulating the production of IL-10 and/or TGF β in a cell or cells of a subject.

The present invention, in a further aspect, relates to use of a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for the manufacture of a medicament for regulating the production of cell surface markers involved in immune responses and antigen recognition such as CD40, I-A/I-E, CD317/BST-2, CD103, CD80, CD86, CD83 and/or Siglec-H and/or the species equivalent in a cell or cells of a subject.

In a further aspect, the present invention relates to use of a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for the manufacture of a medicament for regulating (e.g. downregulating) the expression of one of more Type I IFN genes (such as one or more genes selected from the group consisting of but not limited to IFN- β 1, IFN- β 3, *Ifi202b*, *Ifi203*, *IFI44*, *IFTI*, *MXI*, *OASI*, *OAS2*, *OAS3*, *OASL*, *Irf3* and *Irf4*) in a cell or cells of a subject.

In another aspect, the present invention relates to use of a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for the manufacture of a medicament for regulating (e.g. downregulating) the expression of one or more pro-inflammatory genes (such as one or more genes selected from the group consisting of but not limited to IL1- β , IL4, IL5, IL6, IL8, IL12, IL13, IL17, IL21, IL22, IL23, IL27, IFN γ , CCL2, CCL3, CCL5, CCL20, CXCL5, CXCL10, CXCL12, CXCL13, and TNF- α) in a cell or cells of a subject.

The present invention, in a further aspect, relates to use of a *Roseburia* flagellin, and/or a

polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for the manufacture of a medicament for improving intestinal microbiota in a subject.

The present invention, in another aspect, relates to use of a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for the manufacture of a medicament for regulating appetite in a subject.

The present invention, in a further aspect, relates to use of a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for the manufacture of a medicament for regulating (e.g. downregulating) the expression of the gene encoding cholecystokinin (Cck) and/or the expression of the gene encoding glucagon (Gcg) in a cell or cells of a subject.

The present invention, in another aspect, relates to use of a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for the manufacture of a medicament for improving alimentary canal health in a subject.

In another aspect, the present invention relates to use of a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in medicine.

In a further aspect, the present invention relates to a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said

polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in restoring immunological tolerance.

5

In another aspect, the present invention relates to use of a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for the manufacture of a medicament for restoring immunological tolerance in a subject.

10

The present invention relates to, in another aspect, a method of restoring immunological tolerance in a subject, said method comprising administering a *Roseburia* flagellin and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and wherein the immunological tolerance in a subject is restored.

15

20

In another aspect, the present invention relates to a method of treating a disorder in a subject, said method comprising administering to the subject a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, wherein said disorder is an inflammatory disorder and/or an autoimmune disorder, and wherein *Roseburia hominis* bacteria *per se* are excluded

25

In another aspect, the present invention relates to use of a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, for the manufacture of a medicament for the treatment of a disorder in a subject, wherein said disorder is an inflammatory disorder and/or an autoimmune disorder, and wherein *Roseburia hominis* bacteria *per se* are excluded.

30

35

In another aspect, the present invention relates to a method of stimulating the appetite in a subject, said method comprising administering to the subject a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and wherein the appetite in the subject is stimulated, and wherein *Roseburia hominis* bacteria *per se* are excluded.

In another aspect, the present invention relates to use of a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, for the manufacture of a medicament for stimulating the appetite in a subject, and wherein *Roseburia hominis* bacteria *per se* are excluded.

In another aspect, the present invention relates to a pharmaceutical composition and/or a nutritional supplement comprising *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and a pharmaceutically and/or nutritionally acceptable excipient, carrier or diluent, and wherein a pharmaceutical composition and/or nutritional supplement comprising *Roseburia hominis* bacteria *per se* is excluded.

Throughout this specification the word "comprise", or variations such as "comprises" or "comprising", will be understood to imply the inclusion of a stated element, integer or step, or group of elements, integers or steps, but not the exclusion of any other element, integer or step, or group of elements, integers or steps.

Any discussion of documents, acts, materials, devices, articles or the like which has been included in the present specification is not to be taken as an admission that any or all of these matters form part of the prior art base or were common general knowledge in the field relevant to the present disclosure as it existed before the priority date of each of the appended claims.

FIGURES

Figure 1. Sequence and annotation of *R. hominis* genome. The complete genome sequence of *R. hominis* A2-183 was produced. It is represented by a single 3,592,125-bp chromosome with four ribosomal operons, 66 RNAs and 3,273 predicted proteins. (A) *R. hominis* circular genome map with the location of the PCR experiments (Table S1) indicated in the regions targeted by the primers. The tracks on the genome map, starting at the outer track 0, are: track 0 (blue) - Real-time PCR experiments indicated by numbered tick marks; track 1 (pale blue) - Forward CDS; track 2 (pale blue) - Reverse CDS; track 3 (grey) - rRNA; track 4 (green) - tRNA; track 5 (red) - STS marking regions targeted by Real-time PCR; graph 1 - GC content; graph 2 - GC bias. (B) Functional annotation of the *R. hominis* genome shows that the largest number of genes belong to Carbohydrates , Protein Metabolism and Amino acids and Derivatives.

Figure 2. *R. hominis* responds to the gut environment by up-regulating motility, mobilization and chemotaxis genes. Germfree GF C3H/HeN male mice were given *R. hominis* culture by gavage for 28d, and compared to germfree control animals. At 14d and 28d, *R. hominis*-treated animals (N=5) and control animals (N=4) were sacrificed and ileum,

colon and caecum were collected. (A) Real-time PCR validation of genes involved in conjugation/mobilization transfer. (B) Real-time PCR validation of genes involved in Motility and Chemotaxis. (C) Western blot of *R. hominis* grown *in vitro* in the presence of UV irradiated standard murine chow immuno-stained with affinity-purified Fla1 antibody, Fla2 specific antiserum and anti-DNA gyrase A antibody (lane 1: no diet, lane 2: 0.01g diet/10 mL of *R. hominis* culture, lane 3: 0.02g diet/10 mL, lane 4: 0.05g diet/10 mL, lane 5: 0.1g diet/10 mL, lane 6: 0.2g diet/10 mL, lane 7: 0.5g diet/10 mL, lane 8: 1g diet/10 mL). Electron microscopy (EM) picture of *R. hominis* showing flagella (black arrows). Immunocytochemistry performed on bacteria from luminal contents of colonized mice and from *R. hominis* grown *in vitro* using FlaA1 and FlaA2 specific antiserum. Original magnification x1000. (D) Real-time PCR validation of genes involved in butyrate metabolism. (E) Real-time PCR analysis of *R. hominis* transcripts during *in vitro* exposure to human intestinal epithelial cells. Real-time PCR results are presented as fold change, * $P<0.05$, ** $P<0.01$, *** $P<0.001$.

Figure 3. Identification of transcripts differentially expressed in the murine gut after mono-association with *R. hominis*. (A) Affymetrix microarray analysis of differentially expressed genes in *R. hominis*-colonized mice (N=5) relative to GF (N=4). Bar graphs represent number of genes higher and lower expressed after 14 and 28 days. (B) Heatmap generated from differentially expressed genes with functional significance between GF and *R. hominis*-colonized mice at 14d and 28d. (C) Real-time PCR validation of genes shown to be significantly different between *R. hominis*-colonized and GF mice. Real-time PCR results are presented as fold change, * $P<0.05$, ** $P<0.01$, *** $P<0.001$.

Figure 4. Induction of FoxP3+ Treg cells by *Roseburia hominis*. Flow cytometry analysis of FoxP3+ Treg cells in lamina propria ($P=0.0425$ between control and *R. hominis* treatment) and mesenteric lymph nodes ($P=0.0683$) of conventional C3H/HeN treated for 14 days with *R. hominis*.

Figure 5. Colonic T cell markers are significantly induced by *R. hominis* mono-colonization. (A) Immunofluorescence analysis of ascending colon lamina propria cells with anti-CD3 and anti-FoxP3 in germfree and *R. hominis* mono-colonized C3H/HeN (N=8) and C57Bl/6 (N=3) mice. (B) Immunofluorescence analysis of lamina propria cells labelled with (i) anti-Ly6G, (ii) anti-CD11b, (iii) anti-CD3, and (iv) anti-CD3 and anti-FoxP3 in ascending colon of GF and *R. hominis*-treated C3H/HeN mice. Data expressed as number of positive cells per field of view in GF mice (N=7-8) and *R. hominis*-treated mice (N=8-10). * $P<0.05$. Original magnification x630.

Figure 6. *R. hominis* flagellin RH1 has specific effects on intestinal epithelial cells and murine bone marrow-derived dendritic cells. (A) Heatmap generated from differentially expressed genes in Caco-2 intestinal epithelial cells (N=1) treated with different bacterial flagellins *Salmonella enteritidis* (SE), *E. coli* K12 (EC), RH1 and RH2. (B) Expression of (i) CD40; (ii) I-A/I-E and (iii) CD103 by CD11c+B220+CD317+ Flt3L- derived dendritic cells from conventional C3H/HeN, control (blue) and after 24hrs incubation with recombinant flagellin (SE, K12, RH1, RH2; green) determined by flow cytometry. Histogram represents data from three experiments. (C) Frequencies of recombinant flagellin (SE, K12, RH1, RH2) treated Flt3L- and GM-CSF- derived dendritic cell populations from conventional C3H/HeN gated on CD11c+B220+CD317+ cells and CD11c+CD11b+B220- cells, respectively. Data presented as percentage of total, live and singlet cells, mean $\pm$ SEM from three experiments. (D) Protein expression of cytokines IL-10 and IL-12 was measured by CBA in supernatants from control (unstimulated DCs; N=3) and RH1-treated DCs (N=3) derived from C3H/HeN and C57BL/6. Data is presented as mean $\pm$ SD. *** P <0.001. (E) Quantitative analysis of immunofluorescent labelling of ascending colon lamina propria cells with anti-CD3 and anti-FoxP3 in germfree and *R. hominis* mono-colonized TLR5KO mice (N=2).

Figure 7. *R. hominis* attenuates inflammation in an experimental model of colitis.

Twenty-two female C57BL/6 mice were used to assess the therapeutic effect of *R. hominis* during DSS-induced colitis. Treated mice were dosed daily with 10^9 CFU *R. hominis* for 14 days. From day 8, mice were given DSS in their drinking water for 6 days. The animals were euthanized on day 14 and intestinal tissue sampling was performed. (A) Untreated DSS mice (N=8) had strong elevation of all genes compared to control mice (N=4), while differential gene expression was lower in *R. hominis*-treated animals (N=10). Real-time PCR results are presented as fold change, *** P <0.001. (B) Histopathology tissue scoring presented as mean percentage of fields of view at a given grade. DSS treatment significantly altered all fields of view at grades 0, 2, 3, and 4. *R. hominis* significantly reduced the % fields of view for grade 4 pathology (p =0.02) and increased the % fields of view for grade 0. Data is presented as mean $\pm$ SD. (C) Ascending colon (haematoxylin/eosin stained) of control, DSS-treated and DSS/*R. hominis*-treated IL-10KO animals. Images shown are representative fields of view from each treatment group. Original magnification x100.

Supplementary Figures

Figure S1. *R. hominis* preferentially colonizes the ascending colon of mono-colonized mice. (A) *R. hominis* colonized the ascending colon of the mouse gut with close association of bacteria to the host epithelium, detected by FISH using A2-183 (*R. hominis*-specific probe;

FITC), Eub338 (universal 16S probe; Cy3) and DAPI (nuclei; blue). Overlays of A2-183 + Eub338 and Eub338 + DAPI are also shown. The gamma for the red channel was increased in the Eub338 + DAPI overlay to illustrate the labelled bacteria. Original magnification x630. (B) PCR using *R. hominis*-specific primers showed a strong positive signal in faecal DNA post-colonization, while faeces of GF animals were negative for the presence of any bacteria. (C) Real-time PCR analysis showing colonization levels of *R. hominis*/g faeces. Bacterial DNA isolated from faeces was compared against known standard concentrations of *R. hominis* grown in culture. Similar bacterial levels were detected in all mono-colonized mice, approximating 1×10^{10} bacteria/g faeces. Faeces of GF animals tested negative for the presence of any bacteria.

Figure S2. Gene Ontology analysis performed on genes up-regulated at 28 days in the ascending colon. Gene Ontology (GO) based functional interpretation of the data was performed using DAVID (<http://david.abcc.ncifcrf.gov>). Significantly different transcripts ($P < 0.05$) were allocated into the GO category 'Biological Process' to find significantly enriched GO terms. The GO biological processes for 'actin polymerization' (GO:0030041) and 'negative regulation of I-kappaB kinase/NF-kappaB cascade' (GO:0043124) were affected.

Figure S3. Identification of transcripts differentially expressed in the murine gut after mono-association with *E. coli*. (A) Affymetrix microarray analysis of differentially expressed genes in *E. coli* and *R. hominis*-colonized mice over time. Bar graphs represent number of genes higher and lower expressed after 22 and 28 day, respectively. (B) Heatmap generated from differentially expressed genes with functional significance in *E. coli*-colonized mice at 22d vs. 10d.

Figure S4. Identification of transcripts differentially expressed in TLR5KO mouse gut after mono-association with *R. hominis*. (A) Heatmap generated from differentially expressed genes in ascending colon of *R. hominis*-colonized TLR5KO (N=3) and wildtype mice (N=3) at 28d. (B) Heatmap generated from differentially expressed immune-associated genes in ascending colon of *R. hominis*-colonized TLR5KO and wildtype mice at 28d. (C) Heatmap generated from differentially expressed genes in ileum of *R. hominis*-colonized TLR5KO and wildtype mice at 28d. (D) Heatmap generated from differentially expressed immune-associated genes in ileum of *R. hominis*-colonized TLR5KO and wildtype mice at 28d.

Figure S5. Gene Ontology analysis performed on genes down-regulated at 28 days in

the ascending colon. GO biological processes involved in appetite regulation, such as 'negative regulation of response to food' (GO:0032096), 'negative regulation of appetite' (GO:0032099) and 'regulation of catecholamine secretion' (GO:0050433) were mostly affected.

Figure S6. *R. hominis* colonization influences satiety genes and body weight. Dry body weight and lipid carcass analysis was performed. (A) Dry carcass weights of *R. hominis*-associated mice were significantly heavier compared to GF animals. (B) Carcass lipid analysis showed that total adiposity was also significantly higher in *R. hominis*-treated animals at 14d. Data is presented as mean $\pm$ SD.

Figure 1A Representation of the cloning vector pCR-Blunt II-TOPO used for cloning recombinant flagellins **insoluble after cell lysis**. It allows high-efficiency DNA cloning of blunt-end PCR products. It encodes Kanamycin and Zeocin resistance genes for selection in *E. coli*, and the insert is flanked by multiples restriction sites for excision.

Figure 2A Representation of the expression vector T7-MAT-Tag-FLAG- used for cloning recombinant flagellins insoluble after cell lysis. The multi cloning site (MCS) is flanked by MAT (Metal Affinity Tag) sequence and FLAG peptide (Asp-Tyr-Lys-Asp-Asp-Asp-Lys) sequence, which results in the production of double-tagged flagellin, which can be further purified by affinity columns. This expression vector also encodes a pT7/lac (phage T7 lac operon) promoter for IPTG inducible, high-level expression of MAT-ORF-FLAG recombinant flagellins, an internal lacI gene that represses transcription at basal state in any *E. coli* host, and an AmpR gene for ampicillin selection.

Figure 3A Visualization on a SDS-PAGE of recombinant flagellin cloned with the pCR-Blunt II-TOPO cloning vector and expressed through the pT7-MAT-Tag-FLAG-2 expression vector. Description of lanes: **1**, protein standard (kDa); **2**, RH1.

Figure B1A and B1B show an SDS-Analysis of recombinant flagellins.

- (A) K12 (*Escherichia coli* K12); ER (*Eubacterium rectale* 33656); RI1 (*Roseburia intestinalis* Fla 1);
- RI2 (*Roseburia intestinalis* Fla 2);
- (B) RH1 (*Roseburia hominis* Fla 1); RH2 (*Roseburia hominis* Fla 2); RI3 (*Roseburia intestinalis* Fla 3);
- RI4 (*Roseburia intestinalis* Fla 4).

Figure C1-C4. Multiple sequence alignment confirming flagellin nucleotide sequence and accession number.

Figure D.2 Comparative analysis of CCL20 gene induction by different flagellins. HT-29 and Caco-2 cells were stimulated for 2 hours with a single recombinant flagellin at a concentration of 100 ng/mL. Total RNA was extracted and subjected to real time quantitative PCR analysis for the genes CCL20 and β -actin. Experiments were done in triplicate on three separate occasions. Table D2 indicates significant differences between each treatment calculated with paired t test in HT-29.

Figure D.3 Flagellin-mediated chemokine secretion in HT-29 and Caco-2 cells. Tables D3a, D3b, D3c and D3d indicate significant differences between each treatment calculated with paired t test.

Figure D4, neutralization of TLR5 with an anti-TLR5 specific antibody.

Figure D5A: Frequencies of GM-CSF/IL-4 derived dendritic cells stimulated with recombinant flagellins, data shown as fold change compared to unstimulated GM-CSF/IL-4 derived dendritic cells.

Figure D5B: Frequencies of Flt3L derived dendritic cells stimulated with recombinant flagellins, data shown as fold change compared to unstimulated Flt3L derived dendritic cells.

Figure D 6 A and B Flow cytometry analysis of FoxP3+ Tregs in lamina propria of BOY/J WT and TLR5KO mice treated with *R. hominis*.

DETAILED DESCRIPTION

Flagellin

Flagellin is the principal substituent of bacterial flagellum, and is present in large amounts on nearly all flagellated bacteria. Typically, flagellins are globular proteins which arrange into a hollow cylinder to form the filament in bacterial flagellum.

The diversity of flagellin structural proteins and gene sequences is well recognized and vary according to bacterial species, geographical and clinical and environmental origins. There are thousands of flagellins and related flagellin genes. Some of the important ones in gut

bacteria include the flagellins FLA, FliC, FlgC, FLiE, FlaB, MoA and FliG.

There are several types of FLA (Fla) polypeptides. FlaA1, FlaA2, FlaA3, FlaA4 and FlaB are examples of FLA polypeptides.

Polypeptide FlaA1

The term "polypeptide FlaA1" as used herein refers to the flagellin protein FlaA1. Examples of such polypeptides include the FlaA1 polypeptide of *Roseburia hominis*, *Roseburia cecicola*, *Roseburia faecis*, *Roseburia intestinalis*, and *Roseburia inulinivorans*; the polypeptide sequence shown as SEQ ID NO 2 or SEQ ID No 6; and polypeptides having at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% identity to SEQ ID NO 2 or SEQ ID No 6 or variants, homologues, fragments or derivatives thereof.

SEQ ID NO 2 has the following sequence:

```
MVVQHNLTAMNANRQLGITTGAQAKSSEKLSSGYKINRAADDAAGLTISEKMRSQV
RGLNKASDNAQDGVSLIQVAEGALSETHSILQRMNELATQAANDTNTTSDRTAVQQ
EINQLASEITRIASTTQFNTMNLIDGNFTSKKLQVGSLCGQAITIDISDMSATGLGVSG
LVVSSFSAAGKAMSAAQDAISYVSSMRSKLGALQNRLEHTISHLDNISEHTSSAESRI
RDTDMAEEMVEYSKNNILAQAGQSMLAQANQSTQGVLSLLQ
```

SEQ ID NO 2 is deposited with NCBI under accession number ABI48297.1.

The terms "polypeptide FlaA1" and "FlaA1 polypeptide" are used interchangeably herein.

The terms "FlaA1", "Fla1" and "RH1" may be used interchangeably herein.

In one embodiment, the polypeptide FlaA1 has at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% identity to SEQ ID NO 2 or to variants, homologues, fragments or derivatives thereof. In one aspect, the amino acids at positions 79 to 117 and positions 408 to 439 as well as 411, 412, 420 of SEQ ID NO 2 (or equivalent thereof) are considered important. In one aspect, the amino acids at positions 411, 412, 420 of SEQ ID NO 2 (or equivalent thereof) are alanine (A), glutamine (Q) and serine (S) respectively.

In one embodiment, the polypeptide FlaA1 is the polypeptide shown as SEQ ID NO 2 or SEQ ID No 6.

FlaA1 polypeptides can be derived from certain microorganisms. In one aspect, the FlaA1

polypeptide is derived from a bacterium such as a Firmicute. In a further aspect, the FlaA1 polypeptide is derived from a *Roseburia* spp such as a *Roseburia hominis*, *Roseburia cecicola*, *Roseburia faecis*, *Roseburia intestinalis*, or *Roseburia inulinivorans*.

Roseburia hominis and *Roseburia intestinalis* are recently described commensal gut anaerobes of the phylogenetic cluster XIVa within the Firmicutes phylum, belonging to a dominant group of bacteria in the human gut and are also major butyrate producers. An example of *Roseburia hominis* is the strain deposited under the terms of the Budapest Treaty at National Collections of Industrial, Food and Marine Bacteria (NCIMB) at NCIMB Ltd, Ferguson Building, Craibstone Estate, Bucksburn, Aberdeen, UK, AB21 9YA, on 21 October 2004 by the Rowett Research Institute under the accession number NCIMB 14029^T *Roseburia hominis* A2-183^T (DSM = 16839^T). Another example is the bacterial species is *Roseburia hominis* as described in Duncan, S. H., Aminov, R. I., Scott, K. P., Louis, P., Stanton, T. B., & Flint, H. J. (2006) *Int. J. Syst. Evol. Microbiol.* 56: 2437-2441. An example of *Roseburia intestinalis* is the strain deposited under the accession number NCIMB 13810 *Rosburia intestinalis* L1-82^T (DSM = 14610^T) Another example is the bacterial species is *Roseburia hominis* as described in Duncan, S. H., Aminov, R. I., Scott, K. P., Louis, P., Stanton, T. B., & Flint, H. J. (2006) *Int. J. Syst. Evol. Microbiol.* 56: 2437-2441.

The term "polynucleotide sequence encoding the polypeptide FlaA1" as used herein refers a polynucleotide sequence encoding the flagellin protein FlaA1. Examples of such polynucleotide sequences include the gene *FlaA1* of *R. hominis*, *Roseburia cecicola*, *Roseburia faecis*, *Roseburia intestinalis*, or *Roseburia inulinivorans*; the polynucleotide sequence shown as SEQ ID NO 1 or SEQ ID No 5; polynucleotide sequences encoding the polypeptide shown as SEQ ID NO 2 or SEQ ID No 6; polynucleotides sequences having at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% identity to SEQ ID NO 1 or SEQ ID No 5 or variants, homologues, fragments or derivatives thereof; and polynucleotides sequences encoding the polypeptide shown as SEQ ID NO 2 or SEQ ID No 6 or encoding a polypeptide having at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% identity to the polypeptide shown as SEQ ID NO 2 or SEQ ID No 6 or variants, homologues, fragments or derivatives thereof.

SEQ ID NO 1 has the following sequence:

ATGGTAGTACAGCACAATCTTACAGCAATGAACGCTAACAGACAGTTAGGTATCA
CAACAGGCGCACAGGCTAAGTCTTCTGAGAAGTTATCTTCTGGTTACAAGATCAA
CCGCGCAGCAGATGACGCAGCAGGTCTTACGATTTCCGAGAAGATGAGAAGCC
AGGTTAGAGGCTTAAATAAAGCTTCTGACAACGCACAGGATGGTGTATCCCTTAT

TCAGGTAGCTGAGGGTGCATTAAGTGAGACACACTCCATCTTACAGCGTATGAA
 TGAGTTAGCAACTCAGGCAGCAAACGATACCAATACAACCTCAGACAGAACTGC
 AGTTCAGCAGGAGATCAACCAGTTAGCATCTGAGATCACCAGAATCGCTTCTAC
 AACTCAGTTCAACACAATGAACCTGATCGATGGTAACTTCACAAGTAAGAAGCTT
 CAGGTAGGTTCCCTTTGCGGACAGGCTATCACAATCGATATCTCTGATATGTCTG
 CTACAGGTCTTGGCGTTAGCGGATTAGTAGTATCTTCCTTCTCTGCAGCTGGTAA
 GGCAATGTCTGCAGCTCAGGATGCTATCAGCTACGTATCTTCTATGCGTTCTAAG
 CTGGGTGCATTACAGAACAGACTTGAGCACACAATCTCCCACCTGGACAACATT
 TCTGAGCACACATCTTCTGCAGAGTCTCGTATCCGTGATACAGATATGGCTGAA
 GAGATGGTTGAGTACAGCAAGAACAACATCCTTGCTCAGGCAGGACAGTCTATG
 CTTGCTCAGGCTAACCAGTCTACTCAGGGTGTATTATCCTTATTACAGTAA

SEQ ID NO 1 is deposited with NCBI under accession number DQ789140.1.

In one embodiment, the polynucleotide sequence encoding the polypeptide FlaA1 has at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% identity to the polynucleotide sequence shown as SEQ ID NO 1 or SEQ ID No 5 or to variants, homologues, fragments or derivatives thereof.

In one embodiment, the polynucleotide sequence encoding the polypeptide FlaA1 encodes a polypeptide shown as SEQ ID NO 2 or SEQ ID No 6 or a polypeptide having at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% identity to the polypeptide shown as SEQ ID NO 2 or SEQ ID No 6 or to variants, homologues, fragments or derivatives thereof.

In one embodiment, the polypeptide FlaA1 is a truncated FlaA1 polypeptide. For example, the truncated polypeptide comprises at least 20, 30, 40, 50, 75, 100, 125, 150, 175 or 200 amino acids of polypeptide shown as SEQ ID NO 2 or SEQ ID No 6.

Without wishing to be bound by theory, two essential regions of flagellin protein involved in the recognition and activation of TLR5 are amino acids 79-117 of SEQ ID NO 2 (N-D1 domain) and amino acids 408-439 of SEQ ID NO 2 (C- D1 domain). Without wishing to be bound by theory, amino acid: A411, Q412, S420 are important amino acids.

Examples of truncated FlaA1 polypeptides include: polypeptides which comprise amino acids 79-117 and amino acids 408-439 of SEQ ID NO 2; polypeptides having at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% identity to amino acids 79-117 and amino acids 408-439 of SEQ ID NO 2; polypeptides having at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% identity to amino acids 79-117 and amino acids 408-439 of SEQ ID NO 2 wherein the

amino acid at position 411 (or equivalent thereto) is alanine (A) and/or the amino acid at position 214 is glutamine (Q) and/or the amino acid at position 420 is serine (S); polypeptides comprising amino acids 79-439 of SEQ ID NO 2; polypeptides comprising amino acids 79-439 of SEQ ID NO 2 wherein the amino acid at position 411 (or equivalent thereto) is alanine (A); polypeptides having at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% identity to amino acids 79-439 of SEQ ID NO 2; and polypeptides having at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% identity to amino acids 79-439 of SEQ ID NO 2 wherein the amino acid at position 411 (or equivalent thereto) is alanine (A) and/or the amino acid at position 214 is glutamine (Q) and/or the amino acid at position 420 is serine (S).

In one aspect, the amino acids at positions 411, 412, 420 of SEQ ID NO 2 (or equivalent thereof) in a truncated FlaA1 polypeptide are alanine (A), glutamine (Q) and serine (S) respectively.

In one embodiment, the polynucleotide sequence encoding the polypeptide FlaA1 encodes a truncated FlaA1 polypeptide.

In one embodiment, the polypeptide FlaA1 is a fusion polypeptide. For example, the polypeptide is fused to glutathione *S*-transferase (GST).

Roseburia Genus

Roseburia bacteria are slightly curved rod-shaped cells that are strictly anaerobic and indigenous to the mammalian intestine. The bacteria are butyrate-producing and are actively motile through multiple flagella present along the concave side and in a cluster at one end (Stanton and Savage 1983). Currently, within the *Roseburia* genus, five species have been identified and characterised: *Roseburia cecicola*, *Roseburia faecis*, *Roseburia hominis*, *Roseburia intestinalis*, and *Roseburia inulinivorans* (Stanton and Savage 1983, Duncan *et al* 2006).

Stanton and Savage (1983 - *Roseburia cecicola* gen. nov., sp. nov., a motile, obligately anaerobic bacterium from a mouse cecum. *Int. J. Syst. Bacteriol.*, 1983, **33**, 618-627.) describe *Roseburia cecicola*.

Duncan *et al.* (2006 - Proposal of *Roseburia faecis* sp. nov., *Roseburia hominis* sp. nov. and *Roseburia inulinivorans* sp. nov., based on isolates from human faeces. *Int. J. Syst. Evol. Microbiol.*, 2006, **56**, 2437-2441) describe *Roseburia faecis*.

Duncan *et al.* (2006 - Proposal of *Roseburia faecis* sp. nov., *Roseburia hominis* sp. nov. and *Roseburia inulinivorans* sp. nov., based on isolates from human faeces. *Int. J. Syst. Evol. Microbiol.*, 2006, **56**, 2437-2441) describe *Roseburia hominis*.

Duncan *et al.* (2002 - *Roseburia intestinalis* sp. nov., a novel saccharolytic, butyrate-producing bacterium from human faeces. *Int. J. Syst. Evol. Microbiol.*, 2002, **52**, 1615-1620) describe *Roseburia intestinalis*.

Duncan *et al.* (2006 - Proposal of *Roseburia faecis* sp. nov., *Roseburia hominis* sp. nov. and *Roseburia inulinivorans* sp. nov., based on isolates from human faeces. *Int. J. Syst. Evol. Microbiol.*, 2006, **56**, 2437-2441) describe *Roseburia inulinivorans*.

***Roseburia* Flagellin**

The term “*Roseburia* flagellin” as used herein refers to a flagellin protein derived or derivable from a *Roseburia* bacterium such as *Roseburia. hominis*, *Roseburia cecicola*, *Roseburia faecis*, *Roseburia intestinalis*, and *Roseburia inulinivorans*.

The *Roseburia* flagellin may be flaA1, flaA2, flaA3, flaA4 or combinations thereof.

The terms “FlaA1”, “Fla1”, “flaA1” and “fla1” are used interchangeably herein.

The terms “FlaA2”, “Fla2”, “flaA2” and “fla2” are used interchangeably herein.

The term Fla is used herein to encompass the polypeptides that may be flaA1, flaA2, flaA3, flaA4 or fliC.

In one embodiment, the present invention encompasses the use of at least one *Roseburia* flagellin. For example, the present invention encompasses the use of a combination of at least two, at least three, at least four, or at least five *Roseburia* flagellins.

In some embodiments, the combination of *Roseburia* flagellins comprises flagellins which are derived or derivable from at least two, three, four or five different *Roseburia* species.

Examples of *Roseburia* flagellins include flagellins derived or derivable from *Roseburia* bacterium such as *Roseburia hominis*, *Roseburia cecicola*, *Roseburia faecis*, *Roseburia*

intestinalis, and *Roseburia inulinivorans*. In one embodiment, the flagellin is derived or derivable from *Roseburia hominis*. In another embodiment, the flagellin is derived or derivable from *Roseburia intestinalis*.

5 Examples of *Roseburia* flagellins include polypeptide sequences having at least 75%, 80%, 85%, 90%, 95%, 97%, 98%, or 99% identity to SEQ ID NO 2, SEQ ID NO 4, SEQ ID NO 6, SEQ ID NO 8, SEQ ID NO 10 or SEQ ID NO 12 or variants, homologues, fragments or derivatives thereof.

10 In one embodiment, the *Roseburia* flagellin has the polypeptide sequence shown as SEQ ID NO 2, SEQ ID NO 4, SEQ ID NO 6, SEQ ID NO 8, SEQ ID NO 10 or SEQ ID NO 12.

Examples of *Roseburia hominis* flagellins are *Roseburia hominis* Fla1 and Fla2.

15 An example of *Roseburia hominis* Fla1 is shown as SEQ ID NO 2. *Roseburia hominis* Fla1 is also referred herein as RhFlaA1, RHFlaA1, RhFla1, RHFla1, RH1 or Rh1.

SEQ ID NO 2 has the following sequence:

20 MVVQHNLTA MNANRQLGIT TGAQAKSSEKLSSGYKINRAADDAAGLTISEKMRSQVRGLNK
ASDNAQDGVSLIQVAEGALSETHSILQRMNELATQAANDTNTTSDRTAVQQEINQLASEITRI
ASTTQFNTMNLIDGNFTSKKLQVGS LCGQAITIDISMSATGLGVSGLVVSSFSAAGKAMSA
AQDAISYVSSMR SKLGALQNRLEHTISHLDNISEHTSSAESRIRDTDMAEEMVEYSKNNILAQ
AGQSMLAQANQSTQGVLSLLQ

25 SEQ ID NO 2 is deposited with NCBI under accession number ABI48297.1.

An example of *Roseburia hominis* Fla2 is shown as SEQ ID NO 4. *Roseburia hominis* Fla2 is also referred herein as RhFlaA2, RHFlaA2, RhFla2, RHFla2, Rh2 or RH2.

30 SEQ ID NO 4 has the following sequence:

35 MVVNHNMAAICESRQLRYNVKKMEKSSKKLATGYKLNTANDDAAGLQISETMRHHVKGLNK
ASRNSQDGISMLQTADAALQETQDVLDRMVELTTQAANDINTDSRRRAIQDEL DQLNKEVD
RIAYTTTFNQQYMLAEGTPQAAPGYRIQSGALNGQAIDIHFNASKESLGTDKVNVS SHAK
ASESITMVQDAIEQAALWRDEFGSQQERLEHAVRNTDNTSQNTQSAESGIRDTNMNMEMV
LYSTNRILVHASQSILAQYNDDAKSVLEILK

Examples of *Roseburia intestinalis* flagellins include *Roseburia intestinalis* Fla1, Fla2, Fla3

and Fla4.

An example of *Roseburia intestinalis* Fla1 is shown as SEQ ID No 6. *Roseburia intestinalis* is also referred herein as RiFlaA1, RiFlaA1, RiFla1, RiFla1, Ri1 or RI1.

SEQ ID NO 6 has the following sequence:

MRINYNVSAAIANKHLLGIEDNLSASMERLSSGLKINHSDNPAGMAISNKMKAQIDGLNRAS
QNASDGISVIQIADGALSETTSILQRMRELSVQAASDATMTPADKEAIQKEITSLKDEVDRIST
DTEYNSKTLLDGS�DTRVYTKNATRVDISDHVKAGQYQLSIDTAATQAGPVTANQNYNSTAP
VGASGTMSINGSKVEIEAADTYAEAFKIRNAAETGETTVKIEKNGALSFTAEQYGMSSILEIA
FDDKQLANALGFTADGGNSVVEDPENKGSYVYGGIQNGKVIVPSGTDAEVTLTTPSDGTGF
GDTATVKT DGNKITVTDRAGFEMSFLADAGYTGKLD FVDVDIGTMALHIGANEDQETVRIP
EVSCKS LYIDDADVTTVNGAGRGITQFDDAISKVSEVRSRLGAYQNRLESTVSSLDTFEENM
TGAQSRLTDADMASEMTDYTHQNVLNQAAISVLTQANDLPQ

An example of *Roseburia intestinalis* Fla2 is shown as SEQ ID No 8. *Roseburia intestinalis* Fla2 is also referred herein as RiFlaA2 or RiFlaA2 or Ri2 or RI2.

SEQ ID NO 8 has the following sequence:

MVVNHNMALICESRQLRCNVKNMEKSSKKLATGYKLLGANDDAAGLQISETMRHQTRGLN
KASRNSQDGISMLQTADAALQETQEVLD RMTDLTTQAANDINTDADRRAIQDEIDQLNQEVD
RIAYTTNFNQYILADGTPQARPGYYMIQTGSLAGQGIEIKFVNASKESLGVDKVDVSSHAK
ATESIAVVQNAIEKAASFRDTFGAQQERLEHALRGTDNTSESTQRAESSRRDTNMNMEMV
QYSTNRILVQASQSILAQYNDDAKYVLEMLKQVLQILQ

An example of *Roseburia intestinalis* Fla3 is shown as SEQ ID No 10. *Roseburia intestinalis* Fla3 is also referred herein as RiFla3 or RiFla3 or Ri3 or RI3.

SEQ ID NO 10 has the following sequence:

MVVQHNM TAMNANRMLGVTTSAQAKSSEKLSSGYRINRAGDDAAGLTISEKMRSQIRGLN
KASDNAQDGISLIQVAEGALSETHSILQRMNELATQAANDTNTTADRGAIQDEINQLTSEINRI
SSTTQFNTQNLIDGTFANKNLQVGSICGQRITVSDSMSAGSLNVSANLVKVNTFSAAGEAM
SNIQGAISAISTQRSYLGALQNRLEHTISNLDNISENTQSAESRIRDTDMAEEMVTYSKNNILA
QAGQSMLAQANQSTQGVLSLLQ

An example of *Roseburia intestinalis* Fla4 is shown as SEQ ID No 12. *Roseburia intestinalis* Fla4 is also referred herein as RiFla4 or RiFla4 or Ri4 or RI4.

SEQ ID NO 12 has the following sequence:

MAMVVQHNMSAMNANRNLGVTGTMQAKSSEKLSSGYKINRAADDAAGLSISEKMRSQIR
GLNKASDNAQDGISLIQTAEGALNESHSLQRMRELSVQAANGTETDDDDREAVQNEVSQ
5 QEELTRISSETTEFNTMKLLDGSQSGSTSSTGSGPKFGVVDATLDGALVTSNVKGIKVATAA
ATTTKAGQETAIWAADGKTLTLNLSKNKVYTQDEIDDLIANAKQEDSSATGAPAEVKVSLK
NGIFNADADTTAGTVTAGGVKAVSDEGTVTGFGADTISFTANKYGAEFNDTVFKFKFDKA
AGKEEVETNTAIEIDGANAVTAGEYTIHLAAGKEYTAEDLEDVLKTAGFDFDKLSGNTPDE
PNTLFATSGASTVTDITMGAGTAGAGLGSTDAMWGGAGYDSVSSGAGITLQIGANEGQT
10 MSFSIDDM SARALGVDGNKVDLSTQAGA QKATDTIDAAIKKVS AQRGRMGAIQNRLEHTIS
NLDTA AENTQTAESRIRDTDMAEEMVEYSKNNILAQAGQSMLAQANQSTQGVLSLLQ

The terms “polypeptide FlaA1”, “FlaA1 polypeptide”, “polypeptide Fla1” and “Fla1 polypeptide” are used interchangeably herein. The terms “polypeptide Fla2” and “Fla2 polypeptide” are used interchangeably herein. The terms “polypeptide Fla3” and “Fla3 polypeptide” are used interchangeably herein. The terms “polypeptide Fla4” and “Fla4 polypeptide” are used interchangeably herein.

In one aspect, the *Roseburia* flagellin is selected from the group consisting of Fla1, Fla2, Fla3 and Fla4. In one embodiment, the *Roseburia* flagellin is selected from the group consisting of Fla2, Fla1, Fla4 and combinations thereof. In a further embodiment, the *Roseburia* flagellin is Fla2.

In one aspect, the *Roseburia* flagellin is selected from the group consisting of *Roseburia hominis* Fla1, *Roseburia hominis* Fla2, *Roseburia intestinalis* Fla1, *Roseburia intestinalis* Fla2, *Roseburia intestinalis* Fla3 and *Roseburia intestinalis* Fla4. In one embodiment, the *Roseburia* flagellin is selected from the group consisting of *Roseburia hominis* Fla2, *Roseburia intestinalis* Fla2, *Roseburia intestinalis* Fla1, *Roseburia intestinalis* Fla4 and combinations thereof. In a further embodiment, the *Roseburia* flagellin is selected from the group consisting of *Roseburia hominis* Fla2, *Roseburia intestinalis* Fla2 and combinations thereof. In another embodiment, the *Roseburia* flagellin is *Roseburia intestinalis* Fla2.

In one embodiment, the present invention encompasses the use of at least one polynucleotide sequence encoding a *Roseburia* flagellin. For example, the present invention encompasses the use of a combination of at least two, at least three, at least four, or at least five polynucleotide sequences encoding *Roseburia* flagellins.

In one embodiment, the present invention encompasses the use of a polynucleotide sequence encoding at least one *Roseburia* flagellin. For example, the present invention encompasses the use of a polynucleotide sequence encoding a combination of at least two, at least three, at least four, or at least five polypeptide sequences encoding *Roseburia* flagellins.

The polynucleotide sequence encoding a *Roseburia* flagellin may encode flaA1, flaA2, fla3, fla4 or combinations thereof.

In some embodiments, the combination of polynucleotide sequences encoding *Roseburia* flagellins comprises polynucleotide sequences encoding flagellins which are derived or derivable from at least two, three, four or five different *Roseburia* species.

In some embodiments, the polynucleotide sequence encodes a combination of *Roseburia* flagellins which are derived or derivable from at least two, three, four or five different *Roseburia* species.

In another embodiment, the present invention encompasses the use of a combination of at least one *Roseburia* flagellin and at least one polynucleotide sequence encoding a *Roseburia* flagellin.

Examples of polynucleotide sequence encoding *Roseburia* flagellins include flagellins derived or derivable from *Roseburia* bacterium such as *Roseburia hominis*, *Roseburia cecicola*, *Roseburia faecis*, *Roseburia intestinalis*, and *Roseburia inulinivorans*. In one embodiment, the polynucleotide sequence encodes a *Roseburia* flagellin derived or derivable from *Roseburia hominis*. In another embodiment, the polynucleotide sequence encodes a *Roseburia* flagellin derived or derivable from *Roseburia intestinalis*.

Examples of polynucleotide sequences encoding *Roseburia* flagellin include the polynucleotide sequences encoding a polypeptide which has at least 75% identity to SEQ ID NO 2, SEQ ID NO 4, SEQ ID NO 6, SEQ ID NO 8, SEQ ID NO 10 or SEQ ID NO 12 or variants, homologues, fragments or derivatives thereof and polynucleotide sequences having at least 75% identity to SEQ ID NO 1, SEQ ID NO 3, SEQ ID NO 5, SEQ ID NO 7, SEQ ID NO 9 or SEQ ID NO 11 or variants, homologues, fragments or derivatives thereof.

In one embodiment, the polynucleotide sequence encoding *Roseburia* flagellin has the polynucleotide sequence encoding a polypeptide which has the sequence shown as SEQ ID

NO 2, SEQ ID NO 4, SEQ ID NO 6, SEQ ID NO 8, SEQ ID NO 10 or SEQ ID NO 12.

In one embodiment, the polynucleotide sequence encoding the *Roseburia* flagellin has the polynucleotide sequence shown as SEQ ID NO 1, SEQ ID NO 3, SEQ ID NO 5, SEQ ID NO 7, SEQ ID NO 9 or SEQ ID NO 11.

Examples of polynucleotide sequences encoding *Roseburia hominis* flagellins are polynucleotide sequences encoding *Roseburia hominis* Fla1 and Fla2.

An example of a polynucleotide sequence encoding *Roseburia hominis* Fla1 is shown as SEQ ID NO 1.

SEQ ID No 1 has the following sequence:

ATGGTAGTACAGCACAATCTTACAGCAATGAACGCTAACAGACAGTTAGGTATCACAAC
AGGCGCACAGGCTAAGTCTTCTGAGAAGTTATCTTCTGGTTACAAGATCAACCGCGCAG
CAGATGACGCAGCAGGTCTTACGATTTCCGAGAAGATGAGAAGCCAGGTTAGAGGCTT
AAATAAAGCTTCTGACAACGCACAGGATGGTGTATCCCTTATTCAGGTAGCTGAGGGTG
CATTAAAGTGAGACACACTCCATCTTACAGCGTATGAATGAGTTAGCAACTCAGGCAGCA
AACGATACCAATACAACCTCAGACAGAACTGCAGTTCAGCAGGAGATCAACCAGTTAGC
ATCTGAGATCACCAGAATCGCTTCTACAACCTCAGTTCACACAATGAACCTGATCGATGG
TAACTTCACAAGTAAGAAGCTTCAGGTAGGTTCCCTTTGCGGACAGGCTATCACAATCG
ATATCTCTGATATGTCTGCTACAGGTCTTGGCGTTAGCGGATTAGTAGTATCTTCCTTCT
CTGCAGCTGGTAAGGCAATGTCTGCAGCTCAGGATGCTATCAGCTACGTATCTTCTATG
CGTTCTAAGCTGGGTGCATTACAGAACAGACTTGAGCACACAATCTCCACCTGGACAA
CATTTCTGAGCACACATCTTCTGCAGAGTCTCGTATCCGTGATACAGATATGGCTGAAG
AGATGGTTGAGTACAGCAAGAACAACATCCTTGCTCAGGCAGGACAGTCTATGCTTGCT
CAGGCTAACCAGTCTACTCAGGGTGTATTATCCTTATTACAGTAA

An example of a polynucleotide sequence encoding *Roseburia hominis* Fla2 is shown as SEQ ID NO 3.

SEQ ID No 3 has the following sequence:

ATGGTGGTTAATCATAATATGGCGGCAATCTGTGAGAGCAGGCAGCTGCGCTATAACGT
GAAGAAGATGGAAAAATCTTCCAAAAAGCTTGGCGACAGGGTACAAGCTGAACACAGCAA
ATGATGATGCGGCAGGCTTGCAGATATCAGAGACGATGCGGCATCATGTGAAAGGGCT
GAACAAAGCCTCCCGGAATTCACAGGACGGCATCAGTATGCTGCAGACGGCGGATGCA
GCGCTCCAAGAGACGCAGGATGTTCTCGATCGTATGGTGGAGCTGACGACGCAGGCAG
CCAATGACATCAACACAGACTCGGATCGCAGGGCTATTCAGGATGAGTTGGATCAGCTG

AACAAGGAAGTGGACCGCATCGCCTATACGACGCACTTCAATCAGCAGTATATGTTGGC
GGAGGGAACGCCGAGGCGGCACCGGGATATTACCGCATACAGTCCGGGGCACTGAA
CGGACAGGCGATAGATATCCATTTTGTAAATGCGAGCAAGGAGAGCCTTGGCACAGAC
AAAGTGAATGTATCTTCGCATGCGAAGGCGTCGGAATCCATCACGATGGTTCAGGACGC
5 GATTGAACAGGCGGCGCTCTGGAGAGACGAGTTCGGCAGCCAGCAGGAGCGTCTGGA
ACATGCCGTGCGCAATACGGACAACACATCACAAAATACGCAGAGTGCGGAGTCAGGG
ATCAGAGACACCAACATGAATATGGAGATGGTATTATATTTCGACCAACCGGATTCTGGT
GCATGCATCCCAGAGTATTCTGGCACAGTATAATGATGATGCAAAATCAGTGCTTGAGA
TTTTGAAATAG
10

Examples of polynucleotide sequences encoding *Roseburia intestinalis* flagellins are polynucleotide sequences encoding *Roseburia intestinalis* Fla1, Fla2, Fla3 and Fla4.

An example of a polynucleotide sequence encoding *Roseburia intestinalis* Fla1 is shown as
15 SEQ ID NO 5.

SEQ ID No 5 has the following sequence:

ATGCGTGGCGGAGACAATAGAAGGAGAAACAGAATGAGAATTAATTACAATGTGTCAGC
AGCGATTGCGAATAAACATTTACTTGGGAATTGAGGATAATTTAAGTGCATCGATGGAACG
20 GCTTTCATCGGGACTTAAGATCAACCATTCCAAGGACAATCCGGCAGGAATGGCTATTT
CCAACAAGATGAAAGCACAGATTGATGGTTTAAACCGGGCTTCCCAGAATGCATCGGAT
GGTATTTCTGTTATTCAGATCGCAGATGGTGCGCTGAGTGAAACGACCAGTATTTTACA
GCGTATGAGAGAACTTTCCGTGCAGGCAGCGAGTGATGCAACAATGACACCGGCGGAT
AAAGAAGCAATCCAGAAAGAAATCACTTCATTAAGATGAAGTTGACCGTATTTCTACA
25 GATACAGAGTATAACAGCAAAACACTTTTAGATGGTTCATTAGATACCAGGGTTTACACC
AAAAATGCAACAAGAGTGGACATTTCTGATCATGTGAAAGCAGGACAGTATCAGCTTTC
CATTGATACTGCAGCTACACAGGCCGGACCGGTAACAGCAAATCAGAATTATAATTCCA
CAGCACCGGTCCGTGCGTCCGGAACAATGAGTATTAATGGTTCTAAAGTAGAGATAGAG
GCAGCCGACACCTATGCGGAGGCTTTTGAGAAGATCAGAAATGCAGCAGAGACTGGTG
30 AAACAACCGTTAAGATTGAAAAGAATGGAGCACTTTCATTTACCGCAGAACAGTACGGA
ATGTCAAGCATCTTAGAGATCGCATTNNTGATGATAAGCAGCTTGCTAATGCACTTGGAT
TTACAGCAGACGGAGGAAACAGTGTTGTAGAAGATCCAGAGAATAAAGGCAGCTATGTA
TACGGACAGATTCAGAATGGCAAAGTGATCGTACCTTCCGGTACAGATGCCGAAGTAAC
GCTCACAAAACCGAGTGATGGAACCGGATTTGGTGATACAGCTACGGTAAAAACAGATG
35 GAAATAAGATTACGGTTACAGACAGAGCCGGATTTGAGATGTCATTTCTTGCTGATGCA
GGTTATACGGGTAAGCTGGATTTTGATGTCACGGATATCGGAACGATGGCACTTCATAT
TGGAGCAAATGAGGATCAGGAAACAAGAGTGCGTATTCCGGAGGTTTCCTGCAAGAGC
CTTTACATTGATGATGCAGACGTGACGACTGTAAATGGAGCAGGCAGAGGTATCACACA

GTTTGACGATGCCATTTCAAAGGTCAGTGAAGTGCGTTCAAGACTTGGTGCATACCAGA
ATCGTCTTGAGAGTACGGTATCAAGCCTGGATACGTTTGAAGAAAATATGACAGGAGCC
CAGTCACGACTGACAGATGCGGATATGGCATCGGAAATGACAGATTATACACATCAGAA
TGTATTAAATCAGGCAGCAATCTCTGTTTTGACACAGGCAAACGATCTGCCACAGCAGG
TATTGCAGATTCTGCAGTAA

An example of a polynucleotide sequence encoding *Roseburia intestinalis* Fla2 is shown as SEQ ID NO 7.

SEQ ID No 7 has the following sequence:

ATGGTAGTTAATCATAATATGGCATTGATCTGTGAGAGTAGACAGTTACGATGTAATGTG
AAGAACATGGAGAAGTCTTCAAAAAGCTGGCAACAGGTTATAAATTGCTTGGAGCAAA
TGATGATGCAGCAGGATTACAGATATCAGAAACCATGCGTCATCAGACCAGAGGTCTTA
ACAAAGCATCCAGAAATTCGCAAGATGGAATTAGTATGCTGCAGACAGCAGATGCAGCA
TTACAGGAGACACAGGAAGTGTTGGATCGAATGACGGATCTGACAACACAGGCAGCTA
ATGATATCAATACGGATGCGGATCGTCGTGCAATTCAGGATGAAATCGATCAGTTAAATC
AGGAAGTGGATCGTATTGCATATACGACGAATTTAATCAGCAGTATATATTAGCGGATG
GAACTCCGCAGGCAAGACCAGGATACTATATGATACAGACAGGAAGTCTTGCGGGACA
GGGAATAGAGATTAAGTTTGTTAATGCGAGCAAAGAGAGCTTGGGTGTGGACAAGGTTG
ATGTATCATCGCATGCAAAGCGACAGAATCTATAGCAGTGGTACAGAATGCAATTGAA
AAGGCAGCTTCGTTTAGAGATACATTTGGGGCACAACAGGAGCGGTTAGAACACGCATT
GCGTGGAACGGATAATACATCAGAAAGTACACAGAGGGCAGAATCAAGTAGACGCGAT
ACCAACATGAATATGGAGATGGTACAATATTCTACAAACCGTATTTTAGTACAGGCATCT
CAGAGTATTTTAGCACAGTACAATGATGATGCAAAATATGTGTTGGAAATGTTAAAATAG

An example of a polynucleotide sequence encoding *Roseburia intestinalis* Fla3 is shown as SEQ ID NO 9.

SEQ ID No 9 has the following sequence:

ATGGTAGTACAGCACAAATATGACCGCAATGAATGCGAACAGAATGTTAGGCGTTACAAC
AAGCGCACAGGCAAAATCTTCAGAGAAATTATCTTCTGGTTACAGAATCAACCGTGCAG
GTGATGACGCTGCTGGTTTAAACAATTTCTGAGAAGATGAGAAGCCAGATCCGTGGATTA
AACAAAGCTTCTGACAACGCACAGGATGGTATTTCTTAATCCAGGTTGCTGAGGGTGC
ATTATCTGAGACACATTCTATCTTACAGCGTATGAATGAGTTAGCTACTCAGGCTGCTAA
CGATACCAATACAACCTGCTGATAGAGGAGCTATTCAGGATGAGATCAACCAGTTAACAT
CTGAGATTAACAGAATCTCTTCTACAACCTCAGTTCAATACTCAGAACCTCATCGATGGTA
CATTCGCAAATAAAAACCTTCAGGTTGGTTCTATCTGTGGACAGAGAATTACTGTTTCTA
TCGACAGTATGTCTGCTGGTAGCTTAAATGTATCTGCTAACTTAGTAAAGGTTAACACTT

TCAGTGCAGCAGGTGAAGCAATGTCCAATATTCAGGGTGCTATTTCTGCAATTTCTACAC
AGCGTTCTTACTTAGGAGCTCTTCAGAATCGTCTGGAGCATACAATCTCCAACCTGGACA
ACATTTCTGAGAATACTCAGTCTGCTGAATCTCGTATCCGTGATACAGATATGGCTGAAG
AGATGGTTACTTACAGCAAGAACAATATTCTTGCTCAGGCAGGACAGTCTATGCTTGCTC
5 AGGCTAACCAGTCTACTCAGGGTGTA TTTCTCTGTTACAGTAA

An example of a polynucleotide sequence encoding *Roseburia intestinalis* Fla4 is shown as
SEQ ID NO 11.

10 SEQ ID No 11 has the following sequence:

ATGGCAATGGTAGTACAGCACAAATGTCCGCAATGAATGCGAACAGAAATTTAGGTGT
TACAACAGGAATGCAGGCAAATCATCAGAGAAGTTATCTTCCGGTTACAAGATCAACC
GTGCAGCAGATGATGCAGCAGGACTTTCTATTTCTGAGAAGATGAGAAGCCAGATCCGC
GGTTTAAATAAAGCATCTGACAATGCACAGGATGGTATCTCTTTAATCCAGACCGCTGAG
15 GGAGCATTAATGAGTCCCACTCTATTTTACAGAGAATGAGAGAGTTATCCGTACAGGC
AGCCAACGGTACAGAGACAGATGACGACCGCGAGGCAGTACAGAACGAGGTTTCCAG
TTACAGGAAGAGCTGACAAGAATTTCTGAGACAACAGAGTTCAACACGATGAAGCTGCT
GGATGGTTCTCAGAGTGGAAGTACATCTTCAACCGGGTCAGGTCCGAAGTTTGGTGTTG
TAGATGCAACATTAGACGGTGCACTTGTAACATCTAACGTGAAAGGTATTAAAGTAGCAA
20 CAGCAGCTGCCACAACAACAAGGCAGGTCAGGAGACTGCTATCTGGGCTGCTGATGG
AAAGACATTAAC TTTAAATCTTTGAAAAATAAGGTATATACACAGGACGAAATTGATGAC
TTGATCGCAAATGCAAAACAGGAAGACAGTTCTGCAACGGGTGCACCGGCTGAAGTGA
AAGTATCTTTAAGAATGGTATTTTTAATGCAGATGCAGACACA ACTGCCGGA ACTGTAA
CAGCCGGTGGTGTGAAGGCAGTATCTGATGAAGGAACAGTAACTGGATTTGTTGGTGC
25 AGATACAATTTCA TTTACGGCAAATAAGTATGGAGCAGAGTTCAATGATACTGTATTTAAA
TTCAAATTTGATAAGGCAGCAGGCAAAGAAGAAGTAGAGACAAATACAGCAATTGAAATT
GATGGAGCAAATGCGGTAACAGCAGGTGAATATACAATTCATCTTG CAGCAGGCAAAGA
ATATACGGCAGAAGATTTAGAAGATGTTCTTAAAACGGCAGGATTCGACTTTGATGTTAA
ATTAAGTGGAAATACACCAGATGAGCCAAATACTTTATTTGCAACCAGTGGCGCATCAAC
30 TGTGACTGATATTACAATGGGTGCTGGCACCGCCGGAGCTGGTCTTGGAAGTACAGAT
GCTATGTGGGGGCAAGCTGGTTATGACAGTTATCTTCTGGTGCTGGCATTACCTTG CAG
ATTGGTGCAAATGAAGGTCAGACCATGAGTTTCTCTATCGATGACATGAGTGCAAGAGC
ACTTGCGTAGATGGCAACAAAGTTGATTTAAGCACACAGGCTGGCGCACAGAAAGCAA
CTGATACCATTGATGCAGCAATCAAGAAAGTATCTGCACAGCGTGGTAGAATGGGTGCG
35 ATCCAGAACCGTCTGGAGCACACCATCAGCAACCTTGATACAGCAGCAGAGAATACCCA
GACTGCAGAGTCCCGTATCCGTGATACAGATATGGCAGAAGAGATGGTTGAGTACTCCA
AGAACAACATTCTTG CACAGGCAGGTCAGTCTATGCTTG CACAGGCGAACCAGTCTACA
CAGGGTGTA CTCTCCTTATTACAGTAA

In one aspect, the polynucleotide sequence encodes a *Roseburia* flagellin selected from the group consisting of Fla1, Fla2, Fla3 and Fla4. In one embodiment, the polynucleotide sequence encodes a *Roseburia* flagellin selected from the group consisting of Fla2, Fla1, Fla4 and combinations thereof. In a further embodiment, the polynucleotide sequence encodes the *Roseburia* flagellin Fla2.

In one aspect, the polynucleotide sequence encoding *Roseburia* flagellin is selected from the group consisting of a polynucleotide sequence encoding *Roseburia hominis* Fla1, a polynucleotide sequence encoding *Roseburia hominis* Fla2, a polynucleotide sequence encoding *Roseburia intestinalis* Fla1, a polynucleotide sequence encoding *Roseburia intestinalis* Fla2, a polynucleotide sequence encoding *Roseburia intestinalis* Fla3 and a polynucleotide sequence encoding *Roseburia intestinalis* Fla4. In one embodiment, the polynucleotide sequence encoding *Roseburia* flagellin is selected from the group consisting of a polynucleotide sequence encoding *Roseburia hominis* Fla2, a polynucleotide sequence encoding *Roseburia intestinalis* Fla2, a polynucleotide sequence encoding *Roseburia intestinalis* Fla1, a polynucleotide sequence encoding *Roseburia intestinalis* Fla4 and combinations thereof. In a further embodiment, the polynucleotide sequence encoding *Roseburia* flagellin is selected from the group consisting of a polynucleotide sequence encoding *Roseburia hominis* Fla2, a polynucleotide sequence encoding *Roseburia intestinalis* Fla2 and combinations thereof. In another embodiment, the polynucleotide sequence encoding *Roseburia* flagellin is a polynucleotide sequence encoding *Roseburia intestinalis* Fla2.

The therapeutic effect of flagellins were evaluated by the inventors by functional assays.

In one embodiment, the *Roseburia* flagellin is a truncated *Roseburia* flagellin. For example, the truncated *Roseburia* flagellin comprises at least 20, 30, 40, 50, 75, 100, 125, 150, 175 or 200 amino acids of polypeptide shown as SEQ ID NO 2, SEQ ID NO 4, SEQ ID NO 6, SEQ ID NO 8, SEQ ID NO 10 or SEQ ID NO 12.

In one embodiment, the truncated *Roseburia* flagellin is a truncated *Roseburia hominis* flagellin or a truncated *Roseburia intestinalis* flagellin.

In one embodiment, the polynucleotide sequence encoding the *Roseburia* flagellin encodes a truncated FlaA1, Fla2, Fla3 or Fla4. For example, the polynucleotide sequence encodes a truncated *Roseburia* flagellin comprising at least 20, 30, 40, 50, 75, 100, 125, 150, 175 or

200 amino acids of polypeptide shown as SEQ ID NO 2, SEQ ID NO 4, SEQ ID NO 6, SEQ ID NO 8, SEQ ID NO 10 or SEQ ID NO 12.

5 In one embodiment, the polynucleotide sequence encoding the *Roseburia* flagellin encodes a truncated FlaA1 or Fla3.

In one embodiment, the *Roseburia* flagellin is a fusion polypeptide. For example, the polypeptide is fused to glutathione S-transferase (GST).

Table A – summary of the *Roseburia* flagellin sequences described herein and the bacterial strain from which they may be derived.

Bacteria and strain	Abbreviation	Nucleotide Sequence ID	Nucleotide Sequence	Amino Acid Sequence ID	Amino Acid Sequence
Roseburia hominis A2-183(fla 1)	RhFlaA1 or RHFlaA1 or RhFla1 or RHFla1 or Rh1 or RH1	SEQ ID 1	ATGGTAGTACAGCACAAATCTTACAGCAA TGAACGCTAACACAGACAGTTAGGTATCAC AACAGCGCACAGGCTAAGTCTTCTGA GAAGTTATCTTCTGTTACAAGATCAAC CGCGCAGCAGATGACGCAGCAGGTCTT ACGATTTCCGAGAAGATGAGAAGCCAG GTTAGAGGCTTAAATAAAGCTTCTGACA ACGCACAGGATGGTGATCCCTTATTCA GGTAGCTGAGGGTGCTTAAGTGAGAC ACACTCCATCTTACAGCGTATGAATGAG TTAGCAACTCAGGCAGCAACGATACCA ATACAACCTCAGACAGAACTGCAGTTCA GCAGGAGATCAACCAGTTAGCATCTGA GATCACCGAATCGCTTCTACAACCTCAG TTCAACACAATGAACCTGATCGATGGTA ACTTCACAAGTAAGAAGCTTCAGGTAGG TTCCCTTTGCGGACAGGCTATCACAAATC GATATCTGATATGCTCTGTACAGGTC TTGGCGTTAGCGGATTAGTAGTATCTTC CTTCTGTCAGCTGGTAAGGCAATGCTC GCAGCTCAGGATGCTATCAGCTACGTAT CTTCTATGCGTTCTAAGCTGGGTGCATT ACAGAACAGACTTGAGCACACAATCTCC CACCTGGACAACATTTCTGAGCACACAT CTTCTGCAGAGTCTCGTATCCGTGATAC AGATATGGCTGAAGAGATGGTTGAGTA CAGCAAGAACAACATCCTTGCTCAGGC AGGACAGTCTATGCTTGCTCAGGCTAAC CAGTCTACTCAGGGTGATTATCCTTAT TACAGTAA	SEQ ID 2	MVQHNL TAMNANRQLGITGA QAKSSEKLSSGYKINRAADDA GLTISEKMRSQVRGLNKA SDNA QDGVSLIQVAEGAL SETHSILQRMNELATQA ANDTNTTSDRTAVQ QEINQLASEITRIAS TTQFNTMNLIDGNFTSK KLQVGS LCGQAITIDIS DMSATGLGVSGLVSS FSAAGK AMSAAQDAIS YVSSMR SKLGAL QNRLEHTISHLDNISE HTSSAES RIRDTDMA EEMVEYSKN NILAQA GQSMLAQANQSTQGVLSLLQ
Roseburia hominis A2-183 (fla 2)	RhFlaA2 or RHFlaA2	SEQ ID 3	ATGGTGGTTAATCATAATATGGCGGCAA TCTGTGAGAGCAGGCAGCTGCGCTATA ACGTGAAGAAAGATGGAAAAATCTTCCAA AAAGCTTGCACACAGGGTACAAGCTGAA	SEQ ID 4	MVNVHNMAAICESRQLRYNVKK MEKSSKKLATGYKLN TANDDAAGLQISETMR HHVKG LNKASRNS QDGISMLQTADAALQETQDVLD

	RhFla2 or RHFla2 or Rh2 or RH2		CACAGCAAAATGATGATGCGGCAGGCTT GCAGATATCAGAGACGATGCGGCATCA TGTGAAAGGGCTGAACAAAGCCTCCCG GAATTCACAGGACGGCATCAGTATGCT GCAGACGGCGGATGCAGCGCTCCAAGA GACGAGGATGTTCTCGATCGTATGGT GGAGCTGACGACGCGAGGCGAGCCAATGA CATCAACACAGACTCGGATCGCAGGGC TATTCAGGATGAGTTGGATCAGCTGAAC AAGGAAGTGGACCGCATCGCCTATACG ACGCACCTCAATCAGCAGTATATGTTGG CGGAGGGAACGCGCAGGCGGCACCG GGATATTACCGCATACAGTCCGGGGCA CTGAACGGACAGGCGATAGATATCCATT TTGTAATGCGAGCAAGGAGAGCCTTG GCACAGACAAAGTGAATGTATCTTCGCA TGCGAAGCGCTCGGAATCCATCACGAT GGTTCAGGACGCGATTGAACAGGCGGC GCTCTGGAGAGACGAGTTCGGCAGCCA GCAGGAGCGTCTGGAACATGCCGTGCG CAATACGGACAAACACATCACAAAATACG CAGAGTGGGAGTCAGGGATCAGAGAC ACCAACATGAATATGGAGATGGTATTAT ATTCGACCAACCGGATTCTGGTGCATG CATCCAGAGTATTCCTGGCACAGTAA TGATGATGCAAAATCAGTGCCTTGAGATT TTGAAATAG	SEQ ID 6	MRINYNVSAAIANKHLLGIEDNLS ASMERLSSGLKINHSDNPAGM AISNKMKAQIDGLNRSQNASD GISVIQIADGALSETTSILQRMRE LSVQAAASDA TMTPADKEAIQKEI TSLKDEVDRISDTTEYNSKTLDD GSLDTRVYTKNATRVDISDHVKA GQYQLSIDTAATQAGPVTANQN YNSTAPVGASGTMSNGSKVEIE AADTYAEAFEKIRNAAEGETTV KIEKNGALSFTAEQYGMSSILEIA FDDKQLANALGFTADGGNSVVE
Roseburia intestinalis (fla1)	RiFlaA1 or RiFlaA1 or RiFla1 or RiFla1 or Ri1 or Ri1	SEQ ID 5	ATGCGTGGCGGAGACAATAGAAGGAGA AACAGAAATGAGAAATTAATTACAATGTGT CAGCAGCGGATTGCGAATAAACATTTACT TGGAAATTGAGGATAATTTAAGTGCATCG ATGGAACGGCTTTTCATCGGGACTTAAAGA TCAACCATTCGAAGGACAATCCGGCAG GAATGGCTATTTCCAACAAGATGAAAGC ACAGATTGATGGTTTTAAACCGGGCTTCC CAGAAATGCATCGGATGGTATTTCTGTTA TTCAGATCGCAGATGGTGCCTGAGTG AAACGACCAGTATTTTACAGCGTATGAG AGAACTTTCCGTGCAGGCAGCGAGTGA		

			TGCAACAATGACACCGCGGATAAAGA AGCAATCCAGAAAGAAATCACTTCATTA AAAGATGAAGTTGACCGTATTTCTACAG ATACAGAGTATAACAGCAAAACACTTTT AGATGGTTCATTAGATACCAGGGTTTAC ACCAAAATGCAACAAGAGTGGACATTT CTGATCATGTGAAGCAGGACAGTATCA GCTTCCATTGATACGCGCTACACAG GCCGGACCGGTAACAGCAAA TCAGAA T TATAATCCACAGCACCGGTCGGTGCG TCCGGAACAA TGAGTATTAA TGGTTCTA AAGTAGAGTAGAGGCGCGGACACCT ATCGGAGGCTTTTGAGAAGATCAGAA ATGCAGCAGAGACTGGTGAACAACCG TTAAGATTGAAAAGAA TGAGCACCTTC ATTTACCGCAGAACAGTACGGAATGTCA AGCATCTTAGAGATCGCATTNNTGATGA TAAGCAGCTTGCTAATGCAC TTGGATTT ACAGCAGACGGAGGAACACAGTGTGTA GAAGATCCAGAGAATAAAGGCAGCTAT GTATACGGACAGATTCAGAATGGCAA GTGATCGTACCTTCCGGTACAGATGCC GAAGTAACGCTCACAAAACCGAGTGAT GGAACCGGATTTGGTGATACAGCTACG GTAAAACAGATGGAATAAGATTACGG TTACAGACAGAGCCGGATTTGAGATGTC ATTCCTTGCTGATGCAGGTTATACGGGT AAGCTGGATTTTGATGTCACGGATATCG GAACGATGGCATTCTATATTGGAGCAA TGAGGATCAGGAACAAGAGTGC GTAT TCCGGAGGTTTCC TCGAAGAGCCTTTAC ATTGATGATGCAGACGTGACGACTGTAA ATGGAGCAGGCAGAGGTATCACACAGT TTGACGATGCCATTTCAAAGGTCAGTGA AGTGC GTTCAAAGACTTGGTGCATACCA GAATCGTCTTGAGAGTACGGTATCAAGC CTGGATACGTTTGAAGAAAATATGACAG GAGCCCAGTCACGACTGACAGATGCGG ATATGGCATCGGAAATGACAGATTATAC	DPENKGSYVYGGQIQNGKVIVPS GTDAEVTLTKPSDGTGFGDTAT VKTDGNKITVTD RAGFEMSFLAD AGYTGKLD FVDVTDIGTMALHIGA NEDQETRVRIPEVSCSKSLYIDDA DVTTVNGAGRGITQFDDAISKVS EVR SRLGAYQNRLESTVSSLDTF EENMTGAQSRLTDADMASEM T DYTHQNVNLNQA AISVLTQANDLP Q
--	--	--	--	--

				ACATCAGAATGTATTAAATCAGGCAGCA ATCTCTGTTTTGACACAGGCAACGATC TGCCACAGCAGGTATTGCAGATTCTGCA GTAA			
Roseburia intestinalis (fla2)	L1-82	RiFla2 or RiFla2 or Ri2 or Ri2	SEQ ID 7	ATGGTAGTTAATCATATATGGCATTGA TCTGTGAGAGTAGACAGTTACGATGTAA TGTGAAGAACATGGAGAGTCTTCAAAA AAGCTGGCAACAGGTTATAAATTGCTTG GAGCAATGATGATGCAGCAGGATTAC AGATATCAGAAACCATGCGTCATCAGAC CAGAGGTCTTAACAAAGCATCCAGAAAT TCGCAAGATGGAATTAGTAGCTGCAGA CAGCAGATGCAGCATTACAGGAGACAC AGGAAGTGTGGATCGAATGACGGATC TGACAACACAGGCAGCTAATGATATCAA TACGGATGCGGATCGTCGTGCAATTCA GGATGAAATCGATCAGTTAAATCAGGAA GTGGATCGTATTGCATATACGACGAATT TTAATCAGCAGTATATATTAGCGGATGG AACTCCGAGGCAAGACCAGGATACATA TATGATACAGACAGGAAGTCTTGCGGG ACAGGGAATAGAGATTAAAGTTTGTAAAT GCGAGCAAGAGAGCTTGGTGTGGAC AAGTTGATGTATCATCGCATGCAAAAG CGACAGAATCTATAGCAGTGGTACAGAA TGCAATTGAAAAGGCAGCTTCGTTTGA GATACATTTGGGGCACAAACAGGAGCGG TTAGAACACGCATTGCGTGGAAACGGAT AATACATCAGAAAGTACACAGAGGGCA GAATCAAGTAGACGCGATACCAACATGA ATATGGAGATGGTACAAATATTCTACAAA CCGTATTTTAGTACAGGCATCTCAGAGT ATTTAGCACAGTACAAATGATGATGCAA AATATGTTGGAAATGTTAAATAAG	SEQ ID 8	MVVNHNMALICESQLRCNVKN MEKSSKLATGYKLLGANDDAA GLQISETMRHQTRGLNKA SRNS QDGISMLQTADAALQETQEVLD RMTDLTTQAANDINTDADRRAIQ DEIDLNQEVDRYATTNFNQOY ILADGTPQARPGYYMIQTGSLAG QGIEIKFVNASKESLGVDKVDVS SHAKATESIAVVQNAIEKAAASFR DTFGAQQERLEHALRGTDNTSE STQRAESSRRDTNMNMEMVQY STNRILVQASQSILAQYNDDAKY VLEMLKQVLQILQ	
Roseburia intestinalis (fla3)	L1-82	RiFla3 or RiFla3 or Ri3 or Ri3	SEQ ID 9	ATGGTAGTACAGCACAAATATGACCGCAA TGAATGCGAAACAGAATGTTAGCGGTTAC AACAAAGCGCACAGGCAAAATCTTCAGA GAAATTATCTTCTGTTACAGAAATCAAC CGTGCAGGTGATGACGCTGCTGGTTTA	SEQ ID 10	MVVQHNMTAMNANRMLGVTTT AQAKSSEKSSGYRINRAGDDA AGLTISEKMRSQIRGLNKA SDNA QDGLISLIQVAEGALSETHSILQRM NELATQAANDTNTTADRGAIQDE	

			ACAATTTCTGAGAAAGATGAGAAAGCCAGA TCCGTGATTAAACAAAGCTTCTGACAA CGCACAGGATGGTATTTCCTTAATCCAG GTTGCTGAGGGTGCAATTATCTGAGACA CATTCTATCTTACAGCGTATGAATGAGT TAGCTACTCAGGCTGCTAACGATACCAA TACAACTGCTGATAGAGGAGCTATTACG GATGAGATCAACCAGTTAACATCTGAGA TTAACAGAACTCTTCTACAACTCAGTTC AATACTCAGAACCTCATCGATGGTACAT TCGCAATAAAAACCTTCAGGTGGTTC TATCTGTGGACAGAGAATTACTGTTTCT ATCGACAGTATGCTGCTGGTAGCTTAA ATGTATCTGCTAACTTAGTAAAGGTTAA CACTTTCAGTGCAGCAGGTGAAGCAAT GTCCAATATTCAGGGTGCTATTTCTGCA ATTTCTACACAGCGTTCTTACTTAGGAG CTCTTCAGAAATCGTCTGGAGCATACAAT CTCCAAC TTGGACAACATTTCTGAGAAAT ACTCAGTCTGCTGAATCTCGTATCCGTG ATACAGATATGGCTGAAGAGATGGTTAC TTACAGCAAGAACAAATATTCTTGCTCAG GCAGGACAGTCTATGCTTGCTCAGGCT AACCAGTCTACTCAGGGTGCTACTTTCTC TGTTACAGTAA	INQLTSEINRISSTTQFNTQNLID GTFANKNLQVGSICGQRITVSID SMSAGSLNVSANLVKVNTFSA GEAMSNIQGAISAISTQRSYLGA LQNRLEHTISNLDNISENTQSAE SRIRDTDMAEEMVTYSKNNILAQ AGQSMILAQANQSTQGVL SLLQ	
Roseburia intestinalis (fla4)	RiFla4 or RiFla4 or Ri4 or Ri4	SEQ ID 11	ATGGCAATGGTAGTACAGCACAAACATGT CCGCAATGAATGCGAACAGAAATTTAGG TGTTACAACAGGAATGCAGGCAAAATCA TCAGAGAAGTTATCTCCGGTTACAAGA TCAACCGTGCAGCAGATGATGCAGCAG GACTTCTATTTCTGAGAAGATGAGAAG CCAGATCCGCGGTTTAAATAAAGCATCT GACAATGCACAGGATGGTATCTCTTTAA TCCAGACCGCTGAGGGAGCATTAATG AGTCCCACTCTATTTTACAGAGAAATGAG AGAGTTATCCGTACAGGCAGCCAAACGG TACAGAGACAGATGACGACCGCGAGGC AGTACAGAACGAGGTTTCCAGGTTACAG GAAGAGCTGACAAGAAATTTCTGAGACAA	SEQ ID 12	MAMVVQHNMSAMNANRNLGVT TGMQAKSSEKLSGGYKINRAAD DAAGLSISEKMRSQIRGLNKASD NAQDGLSLIQTAEGALNESHSLQ RMRELSVQAANGTETDDREAV QNEVSQLQEELTRISETTEFNTM KLLDGSQSGSTSSSTGSGPKFGV VDATLDGALVTSNVKGIKVATAA ATTTKAGQETAIAAADGKTLTLN LSKNKVYTQDEIDDLIANAKQED SSATGAPAEVKVSLKNGIFNADA DTTAGTVTAGGVKAVSDEGTVT GFVGADTISFTANKYGAEFNDTV FKFKFDKAAAGKEEVEINTAIEDG

			CAGAGTTCAACACGATGAAGCTGCTGG ATGGTTCTCAGAGTGGAAGTACATCTTC AACCGGGTCAGGTCCGAAGTTTGGTGT TGTAGATGCAACATTAGACGGTGCACTT GTAACATCTAACGTGAAAGGTATTAAG TAGCAACAGCAGCTGCCACAACAACAA AGGAGGTCAGGAGACTGCTATCTGGG CTGCTGATGGAAGACATTAACTTTAA TCTTTCGAAAAA TAAGGTATATACACAG GACGAAATTGATGACTTGATCGCAATG CAAAACAGGAAGACAGTCTCGCAACGG GTGCACCGGCTGAAGTGAAGTATCTTT AAAGAAATGGTATTTTAAATGCAGATGCA GACACAAC TGCCGGAAC TGTAAACAGCC GGTGGTGTGAAGGCAGTATCTGATGAA GGAACAGTAACTGGATTTGTTGGTGCA GATACAATTTCAATTAACGGCAAAATAAGT ATGGAGCAGAGTTCAATGATACTGTATT TAAATTCAAAATTTGATAAGGCAGCAGGC AAAGAAAGAAGTAGAGACAAATACAGCAA TTGAAATTGATGGAGCAAAATGCGGTAAC AGCAGGTGAATATACAAATTCATCTTGCA GCAGGCAAGAATATACGGCAGAAGAT TTAGAAGATGTTCTTAAACGGCAGGAT TCGACTTTGATGTTAAATTAAGTGGAAA TACACCAGATGAGCCAAATACITTTATT GCAACCCAGTGGCGCATCAACTGTGACT GATATTACAATGGGTGCTGGCACCGCC GGAGCTGGTCTTGGAAGTACAGATGCT ATGTGGGGCAAGCTGGTTATGACAGT TATCTTCTGGTGTGGCATTACCTTGCA GATTGGTGCAAAATGAAGTCCAGACCAT GAGTTTCTCTATCGATGACATGAGTGCA AGAGCACTTGGCGTAGATGGCAACAAA GTTGATTTAAGCACACAGGCTGGCGCA CAGAAAGCAACTGATACCATTTGATGCAG CAATCAAGAAAGTATCTGCACAGCGTG GTAGAATGGGTGCGATCCAGAACCGTC TGGAGCACACCATCAGCAACCTTGATAC	ANAVTAGEYTIHLAAGKEYTAED LEDVLKTAGDFDVKLSGNTPDE PNTLFATSGASTVTDITMGAGTA GAGLGSTDAMWQGAGYDSVSS GAGITLQIGANEGQTMFSIDDM SARALGVDGNKVLDLSTQAGAOK ATDTIDAAIKKVSQGRMGAIQ NRLEHTISNLDTAAENTQTAE SRI RDTDMAEEMVEYSKNNILAQAG QSM LAQANQSTQGVLSLLQ
--	--	--	---	---

				AGCAGCAGAGAATACCCAGACTGCAGA GTCCCGTATCCGTGATACAGATATGGCA GAAGAGATGGTTGAGTACTCCAAGAAC AACATTCTTGCACAGGCAGGTCAGTCTA TGCTTGACACAGGCGAACCCAGTCTACAC AGG GTGTACTCTCCTTATTACAGTAA		
--	--	--	--	--	--	--

Table B – summary of the accession numbers for *Roseburia* flagellins described herein.

Bacteria and strain	Abbreviation	Genome or Gene Number	Protein Number (GenBank)	Accession	Protein Number (NCBI Sequence)	Accession Reference
<i>Roseburia hominis</i> A2-183 (fla 1)	RhFlaA1 or RhFlaA1 or Rh1 or RH1	DQ789140.1	ABI48297.1		ABI48297.1	
<i>Roseburia hominis</i> A2-183 (fla 2)	RhFlaA2 or RhFlaA2 or Rh2 or RH2	DQ789141.1	ABI48298.1		ABI48298.1	
<i>Roseburia intestinalis</i> L1-82 (fla1)	RiFlaA1 or RiFlaA1 RiFla1 or RiFla1 or Ri1 or Ri1	ABYJ02000009.1	EEV02820		WP_006855378	
<i>Roseburia intestinalis</i> L1-82 (fla2)	RiFla2 or RiFla2 or Ri2 or Ri2	ABYJ02000032.1	EEV02466		WP_006855745	
<i>Roseburia intestinalis</i> L1-82 (fla3)	RiFla3 or RiFla3 or Ri3 or Ri3	ABYJ02000104.1	EEV00779		WP_006857364	
<i>Roseburia intestinalis</i> L1-82 (fla4)	RiFla4 or RiFla4 or Ri4 or Ri4	ABYJ02000202.1	EEU99488		WP_006858703	

Modulation/regulation

The terms “modulation” and “regulation” may be used interchangeably herein.

In one embodiment, the term “modulation” refers to an increase and/or induction and/or promotion and/or activation. In an alternative embodiment, the term “modulation” refers to a decrease and/or reduction and/or inhibition.

In one embodiment, the term “regulation” refers to an upregulation. In an alternative embodiment, the term “regulation” refers to a downregulation.

In one embodiment, the *Roseburia* flagellin, and/or Fla (such as FlaA1 or FlaA2) polypeptide, and/or the polynucleotide sequence encoding the *Roseburia* flagellin and/or the Fla (such as FlaA1 or FlaA2) polypeptide, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), reduces the inflammation of a tissue or an organ.

For example, inflammation of the alimentary canal or part thereof (such as the intestine) is reduced.

The term “inflammation” as used herein refers to one or more of the following: redness, swelling, pain, tenderness, heat, and disturbed function of a tissue or organ due to an inflammatory process triggered by over-reaction of the immune system.

A reduction in inflammation in a subject can be determined by determining the levels of pro-inflammatory cytokines and chemokines in tissue, serum and/or faecal samples in a subject before, and after, the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. For example, the levels of one or more of the following can be monitored: IL-1, IL-4, IL5, IL6, IL-8, IL-12, IL-13, IL-17, IL-21, IL-22, IL23, TNF α , IFN γ , CXCL1, CXCL10, CCL2, CCL20 serum and faecal calprotectin, SA1009/SA1008 calcium binding proteins, and Type 1 interferons,

CD markers such as CD163, CD14, inflammatory transcription factors such as NF-kB, STAT, and MAPkinases, c-reactive protein (CRP), erythrocyte sedimentation rate (ESR), complement proteins, serum albumin, histological evaluation of target tissues and organs, disease activity indices. Further, in human studies quality of life (QoL) questionnaires can be carried out before, and after, the *Roseburia* flagellin, and/or polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

In one embodiment, the amount of a tissue or organ which is inflamed in a subject is at least 10%, 20%, 30%, 40% or 50% lower when compared to the amount of tissue or organ which is inflamed in a subject before the *Roseburia* flagellin, and/or polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

In one embodiment, the *Roseburia* flagellin, and/or Fla (such as FlaA1 or FlaA2) polypeptide, and/or the polynucleotide sequence encoding the *Roseburia* flagellin and/or Fla (such as FlaA1 or FlaA2) polypeptide, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), reduces the inflammation by epithelial cells of the tissue or the organ.

For example, the epithelial cells are epithelial cells of the alimentary canal or part thereof (such as the intestine).

In one embodiment, the *Roseburia* flagellin, and/or Fla (such as FlaA1 or FlaA2) polypeptide, and/or the polynucleotide sequence encoding the *Roseburia* flagellin and/or the Fla (such as FlaA1 or FlaA2) polypeptide, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), increases the production of T cells in

a subject.

In one embodiment, the T cells are regulatory T cells (also referred to as Tregs) such as regulatory T cells capable of expressing TLR5 (Toll-like receptor 5).

Without wishing to be bound by theory, an increase in Treg numbers will combat the effects of other effector T cells (also referred to as Teffs), such as Th1, Th17 and Th2 which drive inflammation, autoimmunity and allergic/atopic conditions. Hence this property of the *Roseburia* flagellin, and/or the Fla (such as FlaA1 or FlaA2) polypeptide, and/or the polynucleotide sequence encoding the *Roseburia* flagellin and/or the Fla (such as FlaA1 or FlaA2) polypeptide, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), can be exploited to address many diseases where Teff/Treg cell balance is lost, e.g. Crohn's and ulcerative colitis.

In one embodiment, the production of T cells in a subject is increased such that there are at least 10%, 20%, 30%, 40% or 50% more T cells, or greater than 100% more T cells, when compared to the number of T cells in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

The term "immune system" as used herein may refer to the adaptive immune system and/or the innate immune system.

In one aspect, the invention relates to the *Roseburia* flagellin, and/or the Fla (such as FlaA1 or FlaA2) polypeptide, and/or the polynucleotide sequence encoding the *Roseburia* flagellin and/or the Fla (such as FlaA1 or FlaA2) polypeptide, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in regulating the adaptive immune system of a subject.

As used herein, the term “adaptive immune system”, otherwise known as the “specific immune system” refers to highly specialized, systemic cells and processes that eliminate or prevent pathogenic growth. The adaptive immune response provides the vertebrate immune system with the ability to recognize and remember specific pathogens (to generate immunity), and to mount stronger attacks each time the pathogen is encountered.

As used herein, the term “regulating the adaptive immune system” means inducing the activity of the adaptive immune system, and/or promoting immune homeostatic mechanisms by increasing the level of activity relative to the baseline level of activity. Preferably, the adaptive immune system is modulated towards immune regulation (and not immune activation therefore reducing inflammation).

Defects and disorders associated with the adaptive immune system, particularly related to the function of T cells, are associated with many inflammatory and autoimmune diseases. T cell responses associated with Th1, Th2 and Th17 are associated with atopic, inflammatory and autoimmune diseases. Therapies which improve or increase T regulatory (Tregs) cell populations are important in controlling diseases driven by excessive Th1, Th2 and Th17 cell responses.

Another aspect of the invention relates to the *Roseburia* flagellin, and/or the Fla (such as FlaA1 or FlaA2) polypeptide, and/or the polynucleotide sequence encoding the *Roseburia* flagellin and/or the Fla (such as FlaA1 or FlaA2) polypeptide, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in regulating the innate immune system of a subject.

As used herein, the term “innate immune system”, also known as the non-specific immune system, comprises the cells and mechanisms that provide the host with immediate defence against infection by other organisms in a non-specific manner. This means that the cells of the innate system recognize and respond to pathogens in a generic way, but unlike the adaptive immune system, it does not confer long-lasting or protective immunity to the host.

As used herein, the term “regulating the innate immune system” means inducing the activity of the innate immune system, and/or increasing the level of activity relative to the baseline level of activity such that it promotes immune homeostasis.

Loss or dysregulation of the innate immune function, either due to loss of epithelial barrier, innate immune peptides such as defensins, chemokines and cytokines or defective TLR signalling are associated with increased risk of inflammatory diseases, in several body organs including the gut. Such diseases include inflammatory bowel disease.

In one embodiment, the *Roseburia* flagellin, and/or the Fla (such as FlaA1 or FlaA2) polypeptide, and/or the polynucleotide sequence encoding the *Roseburia* flagellin and/or the Fla (such as FlaA1 or FlaA2) polypeptide, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), restores immunological tolerance.

As used herein, the term “immunological tolerance” refers to the process by which the immune system does not attack an antigen such as a self-antigen.

As used herein the term “restoring immunological tolerance” refers to a restoration in immunological tolerance to one or more antigens (such as a self-antigen) in a subject such that the level of immunological tolerance to the antigen is higher when compared to the levels in a subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

In one embodiment, the *Roseburia* flagellin, and/or the Fla (such as FlaA1 or FlaA2) polypeptide, and/or the polynucleotide sequence encoding the *Roseburia* flagellin and/or the Fla (such as FlaA1 or FlaA2) polypeptide, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), activates dendritic cells and/or epithelial cells.

As used herein the term “activates dendritic cells” refers to an up-regulation of one or more cells markers (such as I-A/I-E cell markers, CD80 and CD86 and CD40) and/or an increase in the production of one or more cytokines (such as IL-10 and TGFβ) by cells (such as dendritic cells) in a subject when compared to the levels in a subject before the *Roseburia*

flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

The term "I-A/I-E" as used herein refers to cell markers of MHC class II.

CD40 has an essential role in immunity, and is one of the best characterized of the costimulatory molecules. This receptor, a member of the tumor necrosis factor receptor family, is expressed by professional antigen-presenting cells, such as dendritic cells. CD40 binds its ligand CD40L, which is transiently expressed on T cells and other non-immune cells under inflammatory conditions.

CD40L is an example of a T cell marker. CD3, CD4, CD25, FoxP3, CTLA-4, Ly6g and CD11b are further examples of T cell markers.

CD80 and CD86 are expressed on antigen-presenting cells (such as dendritic cells) and are required for the development and costimulation of T cell responses. The CD28 and CTLA-4 molecules on T cells serve as receptors for the CD80 and CD86 costimulatory antigens.

CD3, CD4, CD25, FoxP3, CTLA-4, Ly6g and CD11b are examples of markers of colonic T regulatory cells.

Without wishing to be bound by theory, the depletion of Ly6g (e.g. Ly6g6c and ly6g6e) increases infection risk, both gut and respiratory tract and is associated with diseases such as neutropenia. Thus, in one embodiment, the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is for use in treating neutropenia.

Another aspect of the invention relates to the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia*

flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), for use in maintaining immune homeostasis in a subject. As used herein "maintaining immune homeostasis" refers to the self-regulation of the body's immune system to maintain oral tolerance or immune stability in response to changing conditions. Oral tolerance refers to the normal immune responses to food and commensal bacteria in a healthy gut. These are lost in coeliac disease and Inflammatory Bowel Diseases such as Crohn's disease and ulcerative colitis. Thus, in one aspect, the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is for use in treating coeliac disease and Inflammatory Bowel Diseases such as Crohn's disease and ulcerative colitis.

In one embodiment, the numbers of a cell marker on the cell(s) of a subject are up-regulated such that there are at least 10%, 20%, 30%, 40% or 50% more of the cell marker on the cell(s), or greater than 100% more of the cell marker on the cell(s), when compared to the number of the cell marker on the cell(s) of the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. In addition or alternatively, the number of cells in a subject which have the cell marker are increased such that there are at least 10%, 20%, 30%, 40% or 50% more cells which have the cell marker, or greater than 100% more cells which have the cell marker, when compared to the number of cell with the cell marker in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

In one aspect, the cells are T cells.

In another aspect, the cells are cells of the alimentary canal (such as cells of the intestine).

In a further aspect, the cells are colonic and/or small intestinal T regulatory cells and may be either CD4 or CD8 positive.

In one aspect, the cell marker is a T cell marker. In another aspect, the cells marker is a colonic T cell marker.

Markers which are type I-A/I-E are examples of a cell marker. CD40 is another example of a cell marker both found on dendritic cells.

In one embodiment, the level of a cytokine in a subject is increased such that the cytokine level is at least 10%, 20%, 30%, 40% or 50% higher, or greater than 100% higher, when compared to the cytokine level in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

Examples of dendritic cells include bone marrow dendritic cells and gut mucosal dendritic cells.

As used herein the term “activates epithelial cells” refers to an increase in the expression of one or more pro-inflammatory genes by epithelial cells in a subject when compared to the expression levels in a subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

The term “pro-inflammatory gene” as used herein refers to a gene which, when expressed, promotes inflammation. Examples of pro-inflammatory genes include genes encoding but not limited to IL1- β , IL4, IL5, IL6, IL8, IL12, IL13, IL17, IL21, IL22, IL23, IL27, IFN γ , CCL2, CCL3, CCL5, CCL20, CXCL5, CXCL10, CXCL12, CXCL13, and TNF- α .

In one embodiment, the *Roseburia* flagellin, and/or the Fla (such as FlaA1 or FlaA2) polypeptide, and/or the polynucleotide sequence encoding the *Roseburia* flagellin and/or the Fla (such as FlaA1 or FlaA2) polypeptide, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), upregulates the production of a cytokine.

The term “upregulates the production of a cytokine” as used herein refers to an increase in the level of a cytokine in a subject compared to the level of the cytokine in a subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

In one embodiment, the level of the cytokine is increased such that the level is at least 10%, 20%, 30%, 40% or 50% higher, or greater than 100% higher, when compared to the level in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

In another embodiment, the *Roseburia* flagellin, and/or the Fla (such as FlaA1 or FlaA2) polypeptide, and/or the polynucleotide sequence encoding the *Roseburia* flagellin and/or the Fla (such as FlaA1 or FlaA2) polypeptide, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), upregulates the production of IL-10

and/or TGF β .

The term "upregulates the production of IL-10" as used herein refers to an increase in the level of IL-10 in a subject compared to the level of IL-10 in a subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

In one embodiment, the level of IL-10 is increased such that the level is at least 10%, 20%, 30%, 40% or 50% higher, or greater than 100% higher, when compared to the level in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

In some aspects, IL-10 is produced by dendritic cells such as bone marrow derived dendritic cells and gut mucosal dendritic cells in particular CD103+ subsets.

In one embodiment, the *Roseburia* flagellin, and/or the Fla polypeptide, and/or the polynucleotide sequence encoding the *Roseburia* flagellin and/or the Fla (such as FlaA1 or FlaA2) polypeptide, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), upregulates the production cell surface markers involved in immune responses and antigen recognition in a cell or cells of a subject.

Examples of cell surface markers involved in immune responses and antigen recognition include CD40, I-A/I-E, CD317/BST-2, CD103, CD80, CD86, CD83 and/or Siglec-H and/or the species equivalent.

Cell surface markers (e.g. CD317/BST-2) may be referred by different names in different species or the cell surface marker may not yet have been identified on the cells of a

particular species. The term “species equivalent” as used herein encompasses these cell surface markers.

The term “upregulates the production CD40” as used herein refers to an increase in the level of CD40 in a subject compared to the level of CD40 in a subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. For example, the number of cells bearing the cell marker CD40 is increased and/or the number of CD40 markers on a cell is increased.

In one embodiment, the numbers of the CD40 cell marker on the cell(s) of a subject are up-regulated such that there are at least 10%, 20%, 30%, 40% or 50% more of the cell marker on the cell(s), or greater than 100% more of the cell marker on the cell(s), when compared to the number of the cell marker on the cell(s) of the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. In addition or alternatively, the number of cells in a subject which have the cell marker CD40 are increased such that there are at least 10%, 20%, 30%, 40% or 50% more cells which have the cell marker, or greater than 100% more cells which have the cell marker, when compared to the number of cell with the cell marker in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

The term “upregulates the production I-A/I-E” as used herein refers to an increase in the level of I-A/I-E in a subject compared to the level of I-A/I-E in a subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide

sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. For example, the number of cells bearing one or more I-A/I-E cell markers is increased and/or the number of I-A/I-E cell markers on a cell is increased.

In one embodiment, the numbers of the I-A/I-E cell markers on the cell(s) of a subject are up-regulated such that there are at least 10%, 20%, 30%, 40% or 50% more of I-A/I-E cell markers on the cell(s), or greater than 100% more of I-A/I-E cell markers on the cell(s), when compared to the number of I-A/I-E cell markers on the cell(s) of the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. In addition or alternatively, the number of cells in a subject which have I-A/I-E cell markers are increased such that there are at least 10%, 20%, 30%, 40% or 50% more cells which have I-A/I-E cell markers, or greater than 100% more cells which have I-A/I-E cell markers, when compared to the number of cell with I-A/I-E cell markers in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

The term "upregulates the production of CD317/BST-2" as used herein refers to an increase in the level of CD317/BST-2 in a subject compared to the level of CD317/BST-2 in a subject before the *Roseburia* flagellin, and/or the polypeptide Fla, and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. For example, the number of cells bearing one or more

CD317/BST-2 cell markers is increased and/or the number of CD317/BST-2 cell markers on a cell is increased.

In one embodiment, the numbers of the CD317/BST-2 cell markers on the cell(s) of a subject are up-regulated such that there are at least 10%, 20%, 30%, 40% or 50% more of CD317/BST-2 cell markers on the cell(s), or greater than 100% more of CD317/BST-2 cell markers on the cell(s), when compared to the number of CD317/BST-2 cell markers on the cell(s) of the subject before the *Roseburia* flagellin, and/or the polypeptide Fla, and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. In addition or alternatively, the number of cells in a subject which have CD317/BST-2 cell markers are increased such that there are at least 10%, 20%, 30%, 40% or 50% more cells which have CD317/BST-2 cell markers, or greater than 100% more cells which have CD317/BST-2 cell markers, when compared to the number of cell with CD317/BST-2 cell markers in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

The term "upregulates the production of CD103" as used herein refers to an increase in the level of CD103 in a subject compared to the level of CD103 in a subject before the *Roseburia* flagellin, and/or the polypeptide Fla, and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. For example, the number of cells bearing one or more CD103 cell markers is increased and/or the number of CD103 cell markers on a cell is increased.

In one embodiment, the numbers of the CD103 cell markers on the cell(s) of a subject are up-regulated such that there are at least 10%, 20%, 30%, 40% or 50% more of CD103 cell markers on the cell(s), or greater than 100% more of CD103 cell markers on the cell(s), when

compared to the number of CD103 cell markers on the cell(s) of the subject before the *Roseburia* flagellin, and/or the polypeptide Fla, and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. In addition or alternatively, the number of cells in a subject which have CD103 cell markers are increased such that there are at least 10%, 20%, 30%, 40% or 50% more cells which have CD103 cell markers, or greater than 100% more cells which have CD103 cell markers, when compared to the number of cell with CD103 cell markers in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

The term "upregulates the production of CD80" as used herein refers to an increase in the level of CD80 in a subject compared to the level of CD80 in a subject before the *Roseburia* flagellin, and/or the polypeptide Fla, and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. For example, the number of cells bearing one or more CD80 cell markers is increased and/or the number of CD80 cell markers on a cell is increased.

In one embodiment, the numbers of the CD80 cell markers on the cell(s) of a subject are up-regulated such that there are at least 10%, 20%, 30%, 40% or 50% more of CD80 cell markers on the cell(s), or greater than 100% more of CD80 cell markers on the cell(s), when compared to the number of CD80 cell markers on the cell(s) of the subject before the *Roseburia* flagellin, and/or the polypeptide Fla, and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. In addition or alternatively, the number of cells in a subject which have CD80 cell

markers are increased such that there are at least 10%, 20%, 30%, 40% or 50% more cells which have CD80 cell markers, or greater than 100% more cells which have CD80 cell markers, when compared to the number of cell with CD80 cell markers in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

The term "upregulates the production of CD86" as used herein refers to an increase in the level of CD86 in a subject compared to the level of CD86 in a subject before the *Roseburia* flagellin, and/or the polypeptide Fla, and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. For example, the number of cells bearing one or more CD86 cell markers is increased and/or the number of CD86 cell markers on a cell is increased.

In one embodiment, the numbers of the CD86 cell markers on the cell(s) of a subject are up-regulated such that there are at least 10%, 20%, 30%, 40% or 50% more of CD86 cell markers on the cell(s), or greater than 100% more of CD86 cell markers on the cell(s), when compared to the number of CD86 cell markers on the cell(s) of the subject before the *Roseburia* flagellin, and/or the polypeptide Fla, and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. In addition or alternatively, the number of cells in a subject which have CD86 cell markers are increased such that there are at least 10%, 20%, 30%, 40% or 50% more cells which have CD86 cell markers, or greater than 100% more cells which have CD86 cell markers, when compared to the number of cell with CD86 cell markers in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla, and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said Fla, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species

Roseburia hominis, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

The term “upregulates the production of CD83” as used herein refers to an increase in the level of CD83 in a subject compared to the level of CD83 in a subject before the *Roseburia* flagellin, and/or the polypeptide Fla, and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. For example, the number of cells bearing one or more CD83 cell markers is increased and/or the number of CD83 cell markers on a cell is increased.

In one embodiment, the numbers of the CD83 cell markers on the cell(s) of a subject are up-regulated such that there are at least 10%, 20%, 30%, 40% or 50% more of CD83 cell markers on the cell(s), or greater than 100% more of CD83 cell markers on the cell(s), when compared to the number of CD83 cell markers on the cell(s) of the subject before the *Roseburia* flagellin, and/or the polypeptide Fla, and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. In addition or alternatively, the number of cells in a subject which have CD83 cell markers are increased such that there are at least 10%, 20%, 30%, 40% or 50% more cells which have CD83 cell markers, or greater than 100% more cells which have CD83 cell markers, when compared to the number of cell with CD83 cell markers in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla, and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said Fla, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

The term “upregulates the production of Siglec-H” as used herein refers to an increase in the level of Siglec-H in a subject compared to the level of Siglec-H in a subject before the *Roseburia* flagellin, and/or the polypeptide Fla, and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla, and/or the vector comprising said

polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. For example, the number of cells bearing one or more Siglec-H cell markers is increased and/or the number of Siglec-H cell markers on a cell is increased.

In one embodiment, the numbers of the Siglec-H cell markers on the cell(s) of a subject are up-regulated such that there are at least 10%, 20%, 30%, 40% or 50% more of Siglec-H cell markers on the cell(s), or greater than 100% more of Siglec-H cell markers on the cell(s), when compared to the number of Siglec-H cell markers on the cell(s) of the subject before the *Roseburia* flagellin, and/or the polypeptide Fla, and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. In addition or alternatively, the number of cells in a subject which have Siglec-H cell markers are increased such that there are at least 10%, 20%, 30%, 40% or 50% more cells which have Siglec-H cell markers, or greater than 100% more cells which have Siglec-H cell markers, when compared to the number of cell with Siglec-H cell markers in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla, and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said Fla, and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

In some aspects, the production of CD40 I-A/I-E, CD317/BST-2, CD80, CD86, CD83 and/or Siglec-H is by dendritic cells (such as tolerogenic CD103⁺ dendritic cells expanded by FLT3L.).

In one embodiment, the expression of one or more Type I IFN genes in a cell or cells of a subject is downregulated.

In one embodiment, the expression level of one or more Type I IFN genes is decreased such that the level is at least 10%, 20%, 30%, 40% or 50% lower when compared to the level in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said

polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

Examples of Type I IFN genes include but not limited to IFN- β 1, IFN- β 3, *Ifi202b*, *Ifi203*, *IFI44*, *IFTI*, *MXI*, *OASI*, *OAS2*, *OAS3*, *OASL*, *Irf3* and *Irf4*.

In one embodiment, the expression of one or more pro-inflammatory genes in a cell or cells of a subject is downregulated.

In one embodiment, the expression level of one or more pro-inflammatory genes is decreased such that the level is at least 10%, 20%, 30%, 40% or 50% lower when compared to the level in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

The term "intestinal microbiota" as used herein refers to microorganisms that live in the digestive tract of the host animals. These microorganisms perform a wide variety of metabolic, structural, protective and other beneficiary functions.

As used herein, the term "improving intestinal microbiota" refers to increasing the number and/or type of microorganisms present in the intestine of a subject (e.g. the host), and/or increasing the activity of said microorganisms in terms of their metabolic, structural, protective and other beneficiary functions. For example, the numbers (i.e. levels) of Clostridium cluster XIVa bacteria are increased and the numbers of *E. coli* are reduced; such an improvement in intestinal microbiota may occur in subjects with inflammatory bowel disease (IBD) once the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), has been administered to the

subject.

In one embodiment, the number of microorganisms present in the intestine of a subject (e.g. the host), is increased such that the number of microorganisms is at least 10%, 20%, 30%, 40% or 50% higher, or greater than 100% higher, when compared to the level in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. In addition, or alternatively, the types of microorganisms present in the intestine of a subject (e.g. the host), are increased such that there are at least 5%, 10%, or 15% more types of microorganisms when compared to the types in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

As used herein, the term “regulating appetite” refers to the ability to modulate (i.e. increase or decrease) the desire for a subject to eat food.

In one embodiment, the appetite in the subject is stimulated (i.e. increased).

Without wishing to be bound by theory, the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), exerts a stimulatory effect on the appetite of a subject by downregulating the expression of genes associated with the suppression of appetite (such as genes encoding satiety hormones). *Agt*, *Cartpt*, *Cck*, *Cxcl12* and *Gcg* are examples of genes associated with regulating appetite and the downregulation of one or more of these genes is associated with the suppression of appetite.

Cck and *Gcg* are examples of satiety hormones.

In one aspect, the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), stimulates the appetite in the subject such that the subject consumes at least 5%, 10%, or 15% more food when compared to the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject. In addition, or alternatively, the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), stimulates the appetite in the subject such that after 1 month from administration the weight of the subject is at least 2%, 5%, or 10% higher when compared to the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

In one embodiment, the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), reduces the level of cholecystokinin

(Cck) and/or glucagon (Gcg) in the blood of a subject.

In one aspect, the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), reduces the level of cholecystokinin (Cck) and/or glucagon (Gcg) in the blood of a subject by at least 5%, 10%, 15% or 20% when compared to the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

In one embodiment, the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or a polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), downregulates the expression of the gene encoding cholecystokinin (Cck) and/or the expression of the gene encoding glucagon (Gcg) in a cell or cells of a subject.

In one aspect, the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), decreases the expression of the gene encoding cholecystokinin (Cck) such that the expression level is at least 5%, 10%, 15% or 20% lower when compared to the expression level in the subject before the *Roseburia* flagellin and/or the polypeptide Fla (such as FlaA1 or FlaA2) and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla (such as FlaA1 or FlaA2) and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial

species *Roseburia intestinalis*), is administered to the subject.

In one aspect, the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), decreases the expression of the gene encoding glucagon (Gcg) such that the expression level is at least 5%, 10%, 15% or 20% lower when compared to the expression level in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

The term “improving alimentary canal health” as used herein refers to reducing the level of inflammation in the alimentary canal or part thereof and/or improving intestinal microbiota.

In one embodiment, the level of inflammation in the alimentary canal is at least 10%, 20%, 30%, 40% or 50% lower when compared to the level of inflammation in the alimentary canal of a subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

In one embodiment, the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or a polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), regulates the expression of at least one gene selected from *Tlr5*, *Tlr1*, *Vnn1*, *Defb37*, *Pla2g*, *Muc16*, *Itln*, *Sprr1a*, *Cldn4*, *Pmp22*, *Crb3*, *Magi3*, *Marveld3*, *Mpp7*, *Defcr20*, *Pcgf2*, *Ltbp4*, *Igsf8* and *Tcfe2a*. Many of these genes

are gut barrier genes and antimicrobials and hence work to reduce invasiveness of gut pathogens and also reduce the numbers of viable pathogens.

In one embodiment, the expression of one or more genes selected from the group consisting of TLR-related genes (e.g. *Tlr5*, *Tlr1*, and *Vnn1*), genes encoding anti-microbial peptides (e.g. *Defb37*, *Pla2g*, *Muc16*, and *Itln*), gut barrier function genes (e.g. *Sprr1a*, *Cldn4*, *Pmp22*, *Crb3*, and *Magi3*), innate immune genes (e.g. *Defcr20*, *Pcgf2*, *Ltbp4*, *Igsf8* and *Tcfe2a*) in a cell or cells of a subject is upregulated.

In one embodiment, the expression of one or more genes selected from the group consisting of TLR-related genes (e.g. *Tlr5*, *Tlr1*, and *Vnn1*), genes encoding anti-microbial peptides (e.g. *Defb37*, *Pla2g*, *Muc16*, and *Itln*), gut barrier function genes (e.g. *Sprr1a*, *Cldn4*, *Pmp22*, *Crb3*, and *Magi3*), innate immune genes (e.g. *Defcr20*, *Pcgf2*, *Ltbp4*, *Igsf8* and *Tcfe2a*) is increased such that the level is at least 10%, 20%, 30%, 40% or 50% higher, or greater than 100% higher, when compared to the level in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

In one embodiment, the expression of one or more genes selected from the group consisting of *Tlr5*, *Tlr1*, *Vnn1*, *Defb37*, *Pla2g*, *Muc16*, *Itln*, *Sprr1a*, *Cldn4*, *Pmp22*, *Crb3*, *Magi3*, *Marveld3*, *Mpp7*, *Defcr20*, *Pcgf2*, *Ltbp4*, *Igsf8* and *Tcfe2a* in a cell or cells of a subject is upregulated.

In one embodiment, the expression level of one or more genes selected from the group consisting of *Tlr5*, *Tlr1*, *Vnn1*, *Defb37*, *Pla2g*, *Muc16*, *Itln*, *Sprr1a*, *Cldn4*, *Pmp22*, *Crb3*, *Magi3*, *Marveld3*, *Mpp7*, *Defcr20*, *Pcgf2*, *Ltbp4*, *Igsf8* and *Tcfe2a* is increased such that the level is at least 10%, 20%, 30%, 40% or 50% higher, or greater than 100% higher, when compared to the level in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered

to the subject.

In one embodiment, the expression of one or more genes selected from the group consisting of genes encoding acetyl-CoA acetyltransferase, 3-hydroxyacyl-CoA dehydrogenase, butyryl-CoA dehydrogenase and phosphoenolpyruvate carboxykinase [ATP] in a cell or cells of a subject is modulated.

In one embodiment, the expression level of one or more genes selected from the group consisting of genes encoding acetyl-CoA acetyltransferase, 3-hydroxyacyl-CoA dehydrogenase, butyryl-CoA dehydrogenase and phosphoenolpyruvate carboxykinase [ATP] is modulated such that the level is at least 10%, 20%, 30%, 40% or 50% higher or lower when compared to the level in the subject before the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is administered to the subject.

In a further aspect, the present invention relates to a transformed microorganism (such as a Firmicutes e.g. a *Roseburia* spp such as *R. hominis* or *R. intestinalis*) in which the expression of a flagellin (e.g. *Roseburia* flagellin) such as FlaA1 or FlaA2 is enhanced compared to the equivalent microorganism prior to transformation, and uses thereof for various therapeutic and nutritional uses as described herein. For example, the transformed microorganism may have been transformed with a nucleotide sequence (such as a promoter) such that the microorganism is capable of upregulating the expression of the gene encoding a flagellin (e.g. *Roseburia* flagellin) such as FlaA1 or FlaA2. In another example, the transformed microorganism may have be transformed with an expression vector comprising a nucleotide sequence encoding a flagellin (e.g. *Roseburia* flagellin) such as FlaA1 or FlaA2 operably linked to a regulatory sequence (such as a promoter) such that the microorganism is capable of overexpressing the gene encoding the flagellin (e.g. *Roseburia* flagellin) such as FlaA1 or FlaA2.

As used herein, the term "expression vector" refers to a DNA construct containing a DNA coding sequence (e.g., gene sequence) that is operably linked to one or more suitable control sequence(s) capable of affecting expression of the coding sequence in a host. Such control sequences include a promoter to effect transcription, an optional operator sequence

to control such transcription, a sequence encoding suitable mRNA ribosome binding sites, and sequences which control termination of transcription and translation. The vector may be a plasmid, a phage particle, or simply a potential genomic insert. Once transformed into a suitable host, the vector may replicate and function independently of the host genome, or may, in some instances, integrate into the genome itself. The plasmid is the most commonly used form of expression vector. However, the description is intended to include such other forms of expression vectors that serve equivalent functions and which are, or become, known in the art.

The term "operably linked" refers to juxtaposition wherein the elements are in an arrangement allowing them to be functionally related. For example, a promoter is operably linked to a coding sequence if it controls the transcription of the coding sequence.

Tissues

In one embodiment, the tissue or organ is the alimentary canal or a section thereof (e.g. the oesophagus, the stomach or the intestine such as the small intestine or the large intestine and colon) or another mucosal sites (such as the nasal passages and the lungs).

In one embodiment, the tissue or organ is the alimentary canal or part thereof.

Examples of parts of the alimentary canal include the oesophagus, the stomach and the intestine (such as the small intestine (e.g. the duodenum, the jejunum and the ileum) and/or the large intestine (e.g. the caecum, ascending colon, transverse colon, descending colon, and sigmoid colon)).

Subject

In one embodiment, the subject is a monogastric animal.

Examples of monogastric animals include poultry, humans, rats, pigs, dogs, cats, horses and rabbits.

In another embodiment, the subject is a mammal such as a monogastric mammal.

Examples of monogastric mammals include omnivores (such as humans, rats, and pigs), carnivores (such as dogs and cats), and herbivores (such as horses and rabbits).

In one embodiment, the subject is a human.

Typically, TLR5 is capable of being expressed in the cells of said subject.

Disorders

Roseburia flagellin, and/or polypeptide Fla (such as FlaA1 or FlaA2), and/or a polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), may be used in treating a disorder in a subject, wherein said disorder is an inflammatory disorder and/or an autoimmune disorder.

In one embodiment, the inflammatory disorder and/or an autoimmune disorder affects the alimentary canal, or a section thereof, of said subject.

In one embodiment, the inflammatory disorder and/or an autoimmune disorder affects a mucosal site of a subject. Examples of mucosal sites include the alimentary canal or a section thereof (e.g. the oesophagus, the stomach or the intestine such as the small intestine or the large intestine and colon), the nasal passages and the lungs.

In one embodiment, the inflammatory disorder and/or an autoimmune disorder is selected from the group consisting of rheumatoid arthritis, psoriasis, multiple sclerosis, type I diabetes, coeliac disease, atopic dermatitis, rhinitis, irritable bowel syndrome (IBS), colitis, inflammatory bowel disorder (IBD), ulcerative colitis, pouchitis, Crohn's disease, functional dyspepsia, atopic diseases, necrotising enterocolitis, and combinations thereof.

In one aspect, the inflammatory disorder is colitis. In a further aspect, the inflammatory disease is Crohn's disease, ulcerative colitis or pouchitis.

In one aspect, the inflammatory disorder and/or an autoimmune disorder affects the intestine.

In one aspect, the intestinal disorder is IBS. The precise pathophysiology of IBS remains to be elucidated. Recent studies have described mucosal inflammation and alterations in intestinal microbiota in IBS patients and a disease correlation with intestinal infections.

In a further aspect, the intestinal disorder is Crohn's disease.

In one aspect, the disorder is an autoimmune disorder.

In one aspect, the autoimmune disorder is selected from the group consisting of ulcerative colitis, pouchitis, rheumatoid arthritis, psoriasis, multiple sclerosis, type I diabetes, allergies (including coeliac disease), atopic dermatitis and rhinitis.

In particular due to its function in restoring immune tolerance, the autoimmune diseases, rheumatoid arthritis, psoriasis, multiple sclerosis, type I diabetes are of particular relevance.

As used herein, the term "medicament" encompasses medicaments for both human and animal usage in human and veterinary medicine. In addition, the term "medicament" as used herein means any substance, which provides a therapeutic and/or beneficial effect. The term "medicament" as used herein is not necessarily limited to substances, which need Marketing Approval, but may include substances which, can be used in cosmetics, nutraceuticals, food (including feeds and beverages for example), probiotic cultures, nutritional supplements and natural remedies. In addition, the term "medicament" as used herein encompasses a product designed for incorporation in animal feed, for example livestock feed and/or pet food.

Prophylactic applications

The *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), according to the invention may also be used in prophylactic applications. In prophylactic applications, the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), according to the invention are administered to a patient susceptible to, or otherwise at risk of, a particular disease in an amount that is sufficient to at least partially reduce the risk of

developing a disease. Such an amount is defined to be "a prophylactic effective dose". The precise amounts depend on a number of specific factors such as the subject's state of health and weight.

Encapsulation

In one embodiment, the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin of the polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is encapsulated.

In a further embodiment, the pharmaceutical composition comprising the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is encapsulated.

In another embodiment, the nutritional supplement comprising the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), is encapsulated.

In a further embodiment, the feedstuff, food product, dietary supplement, or food additive is encapsulated.

The term "encapsulated" as used here refers to a means for protecting the polypeptide, and/or polynucleotide sequence, and/or the vector, and/or the host cell, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), from an incompatible environment by physical separation so that it can be delivered to the target site (e.g. the intestine) without degradation or significant degradation

in order that the polypeptide, and/or polynucleotide sequence, and/or the vector, and/or the host cell, and/or the *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), can have an effect on the target site. An example is an enteric coated capsule.

Even when the objective of the encapsulation is the isolation of the polypeptide, and/or polynucleotide sequence, and/or the vector, and/or the host cell, and/or the *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), from its surroundings, the protective coating or shell must be ruptured at the time of desired action. The rupturing of the protective coating or shell is typically brought about through the application of chemical and physical stimuli such as pressure, enzyme attack, chemical reaction and physical disintegration.

For example, the encapsulation ensures that the polypeptide, and/or polynucleotide sequence, and/or the vector, and/or the host cell, and/or the *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), can be ingested so that the polypeptide, and/or polynucleotide sequence, and/or the vector, and/or the host cell and/or the *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), can be delivered to the intestine (i.e. the target site) in an amount which is effective to produce an effect in the intestine.

Pharmaceutical composition

The pharmaceutical composition may be any pharmaceutical composition. In one aspect, the pharmaceutical composition is to be administered orally, enterally or rectally. For example, the composition may be an edible composition. "Edible" means a material that is approved for human or animal consumption.

The pharmaceutical compositions may be for human or animal usage in human and veterinary medicine.

Examples of such suitable excipients for the various different forms of pharmaceutical compositions described herein may be found in the "Handbook of Pharmaceutical Excipients, 2nd Edition, (1994), Edited by A Wade and PJ Weller.

Acceptable carriers or diluents for therapeutic use are well known in the pharmaceutical art, and are described, for example, in Remington's Pharmaceutical Sciences, Mack Publishing

Co. (A. R. Gennaro edit. 1985).

Examples of suitable carriers include lactose, starch, glucose, methyl cellulose, magnesium stearate, mannitol, sorbitol and the like. Examples of suitable diluents include ethanol, glycerol and water.

The choice of pharmaceutical carrier, excipient or diluent can be selected with regard to the intended route of administration and standard pharmaceutical practice. The pharmaceutical compositions may comprise as, or in addition to, the carrier, excipient or diluent any suitable binder(s), lubricant(s), suspending agent(s), coating agent(s), solubilising agent(s).

Examples of suitable binders include starch, gelatin, natural sugars such as glucose, anhydrous lactose, free-flow lactose, beta-lactose, corn sweeteners, natural and synthetic gums, such as acacia, tragacanth or sodium alginate, carboxymethyl cellulose and polyethylene glycol.

Examples of suitable lubricants include sodium oleate, sodium stearate, magnesium stearate, sodium benzoate, sodium acetate, sodium chloride and the like.

Preservatives, stabilizers, dyes and even flavouring agents may be provided in the pharmaceutical composition. Examples of preservatives include sodium benzoate, sorbic acid and esters of p-hydroxybenzoic acid. Antioxidants and suspending agents may be also used.

Nutritional supplements

Nutritionally acceptable carriers, diluents and excipients include those suitable for human or animal consumption and that are used as standard in the food industry. Typical nutritionally acceptable carriers, diluents and excipients will be familiar to the skilled person in the art.

Feedstuff/products

A further aspect of the invention relates to feedstuffs, food products, dietary supplements and food additives containing *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or a polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said

polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*),.

The terms "feedstuff", "food product" "food additive" and "dietary supplement" as used herein are intended to cover all consumable products that can be solid, jellied or liquid.

Suitable food products may include, for example, functional food products, food compositions, pet food, livestock feed, health foods, feedstuffs and the like. In one aspect, the food product is a health food.

As used herein, the term "functional food product" means food that is capable of providing not only a nutritional effect, but is also capable of delivering a further beneficial effect to the consumer. Accordingly, functional foods are ordinary foods that have components or ingredients (such as those described herein) incorporated into them that impart to the food a specific functional - e.g. medical or physiological benefit - other than a purely nutritional effect.

Examples of specific food products that are applicable to the present invention include milk-based products, ready to eat desserts, powders for re-constitution with, e.g., milk or water, chocolate milk drinks, malt drinks, ready-to-eat dishes, instant dishes or drinks for humans or food compositions representing a complete or a partial diet intended for pets or livestock.

In one aspect, the feedstuff, food product, dietary supplement or food additive according to the present invention are intended for humans, pets or livestock such as monogastric animals. The feedstuff, food product, dietary supplement or food additive may be intended for animals selected from the group consisting of dogs, cats, pigs, horses, or poultry. In a further embodiment, the food product, dietary supplement or food additive is intended for adult species, in particular human adults.

The term "milk-based product" as used herein means any liquid or semi-solid milk or whey based product having a varying fat content. The milk-based product can be, e.g., cow's milk, goat's milk, sheep's milk, skimmed milk, whole milk, milk recombined from powdered milk and whey without any processing, or a processed product, such as yoghurt, curdled milk, curd, sour milk, sour whole milk, butter milk and other sour milk products. Another important group includes milk beverages, such as whey beverages, fermented milks, condensed milks, infant or baby milks; flavoured milks, ice cream; milk-containing food such as sweets.

The feedstuffs, food products, dietary supplements or food additives of the present invention may be - or may be added to - food supplements, also referred to herein as dietary or nutritional supplements or food additives.

The feedstuffs, food products, dietary supplements or food additives according to the invention may also be used in animal nutrition (e.g. in pig nutrition), particularly in the early-weaned period and growing fattening period. The feedstuffs, food products, dietary supplements or food additives are expected to enhance immune function reduce and prevent infectious diseases, beneficially alter the microbiota composition, and improve growth and performance of animals, for example, through increased feed conversion efficiency.

Probiotics or live biotherapeutic product

The *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or a polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), may be used in a probiotic or live biotherapeutic product.

Another aspect of the invention relates to a probiotic composition comprising the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or a polynucleotide sequence encoding said *Roseburia* flagellin and/or said polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*).

As used herein, the term "probiotic" means microbial cell preparations or components of microbial cells with a beneficial effect on the health or well-being of the host. (Salminen S, Ouwehand A. Benno Y. et al "Probiotics: how should they be defined" Trends Food Sci. Technol. 1999;10 107-10).

In one aspect, the probiotic composition is an orally administrable composition of metabolically active, i.e., live and/or or lyophilized, or non-viable heat-killed, irradiated or lysed probiotic bacteria. The probiotic composition may contain other ingredients. The

probiotic composition can be administered orally, i.e., in the form of a tablet, capsule or powder. The probiotic composition may comprise the bacterial species *R. hominis* or *R. intestinalis*. Encapsulated products are favoured for *R. hominis* and *R. intestinalis* as they are anaerobes. Other ingredients (such as vitamin C, for example), may be included as oxygen scavengers and prebiotic substrates (such as these improve the colonisation and survival *in vivo*). Alternatively, the probiotic composition of the invention may be administered orally as a food or nutritional product, such as milk or whey based fermented dairy product, or as a pharmaceutical product.

A suitable daily dose of the probiotic bacteria is from about 1×10^3 to about 1×10^{11} colony forming units (CFU); for example, from about 1×10^7 to about 1×10^{10} CFU; in another example from about 1×10^6 to about 1×10^{10} CFU.

In one aspect, the probiotic composition contains the bacterial species and/or cellular components thereof, as active ingredients, in an amount of from about 1×10^6 to about 1×10^{11} CFU/g, respect to the weight of the composition; for example, from about 1×10^8 to about 1×10^{10} CFU/g. The dose may be of 1g, 3g, 5g, and 10g.

Typically, a probiotic is optionally combined with at least one suitable prebiotic compound. A prebiotic is usually a non-digestible carbohydrate such as an oligo- or polysaccharide, or a sugar alcohol, which is not degraded or absorbed in the upper digestive tract. Known prebiotics include commercial products such as inulin and transgalacto-oligosaccharides.

In one aspect, the probiotic composition of the present invention includes a prebiotic in an amount of from about 1 to about 30% by weight, respect to the total weight composition, (e.g. from 5 to 20% by weight). Carbohydrates may be selected from the group consisting of: fructo- oligosaccharides (or FOS), short-chain fructo-oligosaccharides, inulin, isomalt-oligosaccharides, pectins, xylo-oligosaccharides (or XOS), chitosan-oligosaccharides (or COS), beta-glucans, arable gum modified and resistant starches, polydextrose, D-tagatose, acacia fibers, carob, oats, and citrus fibers. In one aspect, the prebiotics are the short-chain fructo-oligosaccharides (for simplicity shown herein below as FOSs-c.c); said FOSs-c.c. are not digestible carbohydrates, generally obtained by the conversion of the beet sugar and including a saccharose molecule to which three glucose molecules are bonded.

Administration

The pharmaceutical compositions, the nutritional supplements, feedstuffs, food products,

dietary supplements or food additives of the present invention may be adapted for oral, rectal, vaginal, parenteral, intramuscular, intraperitoneal, intraarterial, intrathecal, intrabronchial, subcutaneous, intradermal, intravenous, nasal, buccal or sublingual routes of administration.

In one aspect, the pharmaceutical compositions, the nutritional supplements, feedstuffs, food products, dietary supplements or food additives of the present invention are adapted for oral, rectal, vaginal, parenteral, nasal, buccal or sublingual routes of administration.

In a further aspect, the pharmaceutical compositions, the nutritional supplements, feedstuffs, food products, dietary supplements or food additives of the present invention are adapted for oral administration.

For oral administration, particular use is made of compressed tablets, pills, tablets, gellules, drops, and capsules.

Other forms of administration comprise solutions or emulsions which may be injected intravenously, intraarterially, intrathecally, subcutaneously, intradermally, intraperitoneally or intramuscularly, and which are prepared from sterile or sterilisable solutions. The pharmaceutical compositions of the present invention may also be in form of suppositories, pessaries, suspensions, emulsions, lotions, ointments, creams, gels, sprays, solutions or dusting powders.

An alternative means of transdermal administration is by use of a skin patch. For example, the active ingredient can be incorporated into a cream consisting of an aqueous emulsion of polyethylene glycols or liquid paraffin. In another example, the active ingredient can also be incorporated into an ointment consisting of a white wax or white soft paraffin base together with such stabilisers and preservatives as may be required.

Pharmaceutical compositions, the nutritional supplements, feedstuffs, food products, dietary supplements or food additives may be formulated in unit dosage form, i.e., in the form of discrete portions containing a unit dose, or a multiple or sub-unit of a unit dose.

Dosage

A person of ordinary skill in the art can easily determine an appropriate dose of the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or a

polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), to administer to a subject without undue experimentation. Typically, a physician will determine the actual dosage which will be most suitable for an individual patient and it will depend on a variety of factors including the activity of the specific bacterial strain employed, the metabolic stability and length of action of that strain, the age, body weight, general health, sex, diet, mode and time of administration, rate of excretion, drug combination, the severity of the particular condition, and the individual undergoing therapy. The dosages disclosed herein are exemplary of the average case. There can of course be individual instances where higher or lower dosage ranges are merited, and such are within the scope of this invention.

Combinations

In one aspect, the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or a polynucleotide sequence encoding said *Roseburia* flagellin and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), are administered in combination with at least one or two, or three or four or five other active agents. In such cases, the *Roseburia* flagellin, and/or the polypeptide Fla (such as FlaA1 or FlaA2), and/or a polynucleotide sequence encoding said *Roseburia* flagellin and/or polypeptide Fla (such as FlaA1 or FlaA2), and/or the vector comprising said polynucleotide sequence, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, and/or *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), may be administered consecutively, simultaneously or sequentially with the one or more other active agents.

The at least one or two, or three or four or five other active agents may be selected from the group consisting of: *Roseburia* flagellins, the polypeptide Fla (such as FlaA1 or FlaA2), polynucleotide sequence(s) encoding said *Roseburia* flagellin, polynucleotide sequence(s) encoding said polypeptide Fla (such as FlaA1 or FlaA2), vector(s) comprising said polynucleotide sequence(s), host cell(s) comprising said vector(s), host cell(s) comprising said polynucleotide sequence(s), and microorganisms (e.g. *Roseburia* such as *R. hominis*

and/or *R. intestinalis*).

The at least one or two, or three or four or five other active agents may be a microorganism (e.g. a *Roseburia* such as *R. hominis* and/or *R. intestinalis*). Examples of suitable microorganisms include: *Roseburia hominis* A2-183 and *Roseburia intestinalis* L1-82.

Sequence Identity or Sequence Homology

The terms "polypeptide", "polypeptide sequence", "protein" and "amino acid sequence" are used interchangeably herein.

The terms "polynucleotide sequence" and "nucleotide sequence" are used interchangeably herein.

The present invention also encompasses the use of sequences having a degree of sequence identity or sequence homology with amino acid sequence(s) of a polypeptide described herein (e.g. variants, homologues and derivatives) or of any nucleotide sequence encoding such a polypeptide (hereinafter referred to as a "homologous sequence(s)"). Here, the term "homologue" means an entity having a certain homology with the subject amino acid sequences and the subject nucleotide sequences. Here, the term "homology" can be equated with "identity".

In the present context, a homologous sequence is taken to include an amino acid or a nucleotide sequence which may be at least 50, 60, 70, 75, 80, 85 or 90% identical, in some embodiments at least 95, 96, 97, 98 or 99% identical to the subject sequence. Although homology can also be considered in terms of similarity (i.e. amino acid residues having similar chemical properties/functions), in the context of the present invention it is preferred to express homology in terms of sequence identity.

In some embodiments, a homologous sequence is taken to include an amino acid sequence or nucleotide sequence which has one or several additions, deletions and/or substitutions compared with the subject sequence.

In some embodiments, the present invention relates to the use of a protein whose amino acid sequence is represented herein or a protein derived from this (parent) protein by substitution, deletion or addition of one or several amino acids, such as 2, 3, 4, 5, 6, 7, 8, 9 amino acids,

or more amino acids, such as 10 or more than 10 amino acids in the amino acid sequence of the parent protein and having the activity of the parent protein.

In some embodiments, the present invention relates to the use of a nucleic acid sequence (or gene) encoding a protein whose amino acid sequence is represented herein or encoding a protein derived from this (parent) protein by substitution, deletion or addition of one or several amino acids, such as 2, 3, 4, 5, 6, 7, 8, 9 amino acids, or more amino acids, such as 10 or more than 10 amino acids in the amino acid sequence of the parent protein and having the activity of the parent protein.

In the present context, a homologous sequence is taken to include a nucleotide sequence which may be at least 50, 60, 70, 75, 85 or 90% identical, in some embodiments at least 95, 96, 97, 98 or 99% identical to a nucleotide sequence encoding a polypeptide described herein (the subject sequence). Typically, the homologues will comprise the same or equivalent sequences that code for the domain(s) etc. as the subject sequence. Although homology can also be considered in terms of similarity (i.e. amino acid residues having similar chemical properties/functions), in the context of the present invention it is preferred to express homology in terms of sequence identity.

The homologous amino acid sequence and/or nucleotide sequence may provide and/or encode a polypeptide which retains the functional activity and/or enhances the activity of the polypeptide.

In some aspects, an amino acid sequence as described herein has at least 50, 60, 70, 75, 80, 85 or 90% identity, in some embodiments at least 95, 96, 97, 98 or 99% identity to the subject sequence.

In some aspects, a nucleotide sequence as described herein has at least 50, 60, 70, 75, 80, 85 or 90% identity, in some embodiments at least 95, 96, 97, 98 or 99% identity to the subject sequence.

Homology comparisons can be conducted by eye, or more usually, with the aid of readily available sequence comparison programs. These commercially available computer programs can calculate % homology between two or more sequences.

% homology may be calculated over contiguous sequences, i.e. one sequence is aligned with the other sequence and each amino acid in one sequence is directly compared with the

corresponding amino acid in the other sequence, one residue at a time. This is called an “ungapped” alignment. Typically, such ungapped alignments are performed only over a relatively short number of residues.

Although this is a very simple and consistent method, it fails to take into consideration that, for example, in an otherwise identical pair of sequences, one insertion or deletion will cause the following amino acid residues to be put out of alignment, thus potentially resulting in a large reduction in % homology when a global alignment is performed. Consequently, most sequence comparison methods are designed to produce optimal alignments that take into consideration possible insertions and deletions without penalizing unduly the overall homology score. This is achieved by inserting “gaps” in the sequence alignment to try to maximize local homology.

However, these more complex methods assign “gap penalties” to each gap that occurs in the alignment so that, for the same number of identical amino acids, a sequence alignment with as few gaps as possible - reflecting higher relatedness between the two compared sequences - will achieve a higher score than one with many gaps. “Affine gap costs” are typically used that charge a relatively high cost for the existence of a gap and a smaller penalty for each subsequent residue in the gap. This is the most commonly used gap scoring system. High gap penalties will of course produce optimized alignments with fewer gaps. Most alignment programs allow the gap penalties to be modified. Typically the default values are used when using such software for sequence comparisons.

Calculation of maximum % homology therefore firstly requires the production of an optimal alignment, taking into consideration gap penalties. A suitable computer program for carrying out such an alignment is the Vector NTI (Invitrogen Corp.). Examples of software that can perform sequence comparisons include, but are not limited to, the BLAST package (see Ausubel *et al* 1999 Short Protocols in Molecular Biology, 4th Ed - Chapter 18), BLAST 2 (see FEMS Microbiol Lett 1999 174(2): 247-50; FEMS Microbiol Lett 1999 177(1): 187-8 and tatiana@ncbi.nlm.nih.gov), FASTA (Altschul *et al* 1990 J. Mol. Biol. 403-410) and AlignX for example. At least BLAST, BLAST 2 and FASTA are available for offline and online searching (see Ausubel *et al* 1999, pages 7-58 to 7-60).

Although the final % homology can be measured in terms of identity, the alignment process itself is typically not based on an all-or-nothing pair comparison. Instead, a scaled similarity score matrix is generally used that assigns scores to each pairwise comparison based on chemical similarity or evolutionary distance. An example of such a matrix commonly used is

the BLOSUM62 matrix - the default matrix for the BLAST suite of programs. Vector NTI programs generally use either the public default values or a custom symbol comparison table if supplied (see user manual for further details). For some applications, it is preferred to use the default values for the Vector NTI package.

Alternatively, percentage homologies may be calculated using the multiple alignment feature in Vector NTI (Invitrogen Corp.), based on an algorithm, analogous to CLUSTAL (Higgins DG & Sharp PM (1988), Gene 73(1), 237-244).

Once the software has produced an optimal alignment, it is possible to calculate % homology, for example % sequence identity. The software typically does this as part of the sequence comparison and generates a numerical result.

Should Gap Penalties be used when determining sequence identity, then the following parameters can be used for pairwise alignment for example:

FOR BLAST	
GAP OPEN	0
GAP EXTENSION	0

FOR CLUSTAL	DNA	PROTEIN	
WORD SIZE	2	1	K triple
GAP PENALTY	15	10	
GAP EXTENSION	6.66	0.1	

In one embodiment, CLUSTAL may be used with the gap penalty and gap extension set as defined above.

In one embodiment, the degree of identity with regard to a nucleotide sequence is determined over at least 20 contiguous nucleotides, for example over at least 30 contiguous nucleotides, for example over at least 40 contiguous nucleotides, for example over at least 50 contiguous nucleotides, for example over at least 60 contiguous nucleotides, for example over at least 100 contiguous nucleotides, for example over at least 200 contiguous nucleotides, for example over at least 300 contiguous nucleotides.

In one embodiment, the degree of identity with regard to a nucleotide sequence may be determined over the whole sequence.

Recombinant

In one aspect the *Roseburia* flagellin polypeptide and/or polynucleotide sequence for use in the present invention is a recombinant sequence – i.e. a sequence that has been prepared using recombinant DNA techniques.

These recombinant DNA techniques are within the capabilities of a person of ordinary skill in the art. Such techniques are explained in the literature, for example, J. Sambrook, E. F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Second Edition, Books 1-3, Cold Spring Harbor Laboratory Press.

Synthetic

In one aspect the *Roseburia* flagellin polynucleotide sequence for use in the present invention is a synthetic sequence – i.e. a sequence that has been prepared by *in vitro* chemical or enzymatic synthesis. It includes, but is not limited to, sequences made with optimal codon usage for host organisms - such as the methylotrophic yeasts *Pichia* and *Hansenula*.

Expression of enzymes

The nucleotide sequence for use in the present invention may be incorporated into a recombinant replicable vector. The vector may be used to replicate and express the nucleotide sequence, in protein form, in and/or from a compatible host cell.

Expression may be controlled using control sequences e.g. regulatory sequences.

The protein produced by a host recombinant cell by expression of the nucleotide sequence may be secreted or may be contained intracellularly depending on the sequence and/or the vector used. The coding sequences may be designed with signal sequences which direct secretion of the substance coding sequences through a particular prokaryotic or eukaryotic cell membrane.

Expression vector

In one aspect, the present invention relates to a vector (such as an expression vector) comprising at least one polynucleotide sequence encoding at least one *Roseburia* flagellin.

The term "expression vector" means a construct capable of *in vivo* or *in vitro* expression.

In one embodiment, the expression vector is incorporated into the genome of a suitable host organism. The term "incorporated" in one aspect covers stable incorporation into the genome.

The nucleotide sequence of the present description may be present in a vector in which the nucleotide sequence is operably linked to regulatory sequences capable of providing for the expression of the nucleotide sequence by a suitable host cell or host organism.

The vectors for use in the present invention may be transformed into a suitable host cell or host organism as described herein to provide for expression of a polypeptide of the present description.

The choice of vector e.g. a plasmid, cosmid, or phage vector will often depend on the host cell into which it is to be introduced.

The vectors for use in the present invention may contain one or more selectable marker genes- such as a gene, which confers antibiotic resistance e.g. ampicillin, kanamycin, chloramphenicol or tetracyclin resistance. Alternatively, the selection may be accomplished by co-transformation (as described in WO91/17243).

Vectors may be used *in vitro*, for example for the production of RNA or used to transfect, transform, transduce or infect a host cell.

Thus, in a further embodiment, the description provides a method of making nucleotide sequences of the present description by introducing a nucleotide sequence of the present description into a replicable vector, introducing the vector into a compatible host cell, and growing the host cell under conditions which bring about replication of the vector.

The vector may further comprise a nucleotide sequence enabling the vector to replicate in the host cell in question. Examples of such sequences are the origins of replication of plasmids pUC19, pACYC177, pUB110, pE194, pAMB1 and pIJ702.

An expression vector may comprise at least two, three, four or five polynucleotide sequences encoding *Roseburia* flagellins.

Examples of expression vectors include pGEX-6P-1, pCR-Blunt II-TOPO and T7-MAT-Tag-FLAG-.

The expression vector pGEX-6P-1 may be used for cloning recombinant flagellins. The expression vector pGEX-6P-1 comprises a tac promoter for chemically inducible, high-level expression of GST-tagged recombinant proteins, an internal lacI_q gene for use in any *E. coli* host, an Amp^R gene for ampicillin selection and a PreScission Protease site for cleaving, if desired, the protein from the fusion product .

The cloning vector pCR-Blunt II-TOPO may be used for cloning recombinant flagellins, in particular those which are insoluble after cell lysis. Typically this vector allows high-efficiency DNA cloning of blunt-end PCR products. The vector comprises Kanamycin and Zeocin resistance genes for selection in *E. coli*, and the insert is flanked by multiples restriction sites for excision.

Expression vector T7-MAT-Tag-FLAG- may be used for cloning recombinant flagellins, in particular those which insoluble after cell lysis. The multi cloning site (MCS) is flanked by MAT (Metal Affinity Tag) sequence and FLAG peptide (Asp-Tyr-Lys-Asp-Asp-Asp-Lys) sequence, which results in the production of double-tagged flagellin, which can be further purified by affinity columns. This expression vector also comprises a pT7/lac (phage T7 lac operon) promoter for IPTG inducible, high-level expression of MAT-ORF-FLAG recombinant flagellins, an internal lacI gene that represses transcription at basal state in any *E. coli* host, and an Amp^R gene for ampicillin selection.

Regulatory sequences

In some applications, the nucleotide sequence for use in the present invention is operably linked to a regulatory sequence which is capable of providing for the expression of the nucleotide sequence, such as by the chosen host cell. By way of example, the present invention covers a vector comprising the nucleotide sequence of the present invention operably linked to such a regulatory sequence, i.e. the vector is an expression vector.

The term "operably linked" refers to a juxtaposition wherein the components described are in a relationship permitting them to function in their intended manner. A regulatory sequence "operably linked" to a coding sequence is ligated in such a way that expression of the coding sequence is achieved under condition compatible with the control sequences.

The term "regulatory sequences" includes promoters and enhancers and other expression regulation signals.

The term "promoter" is used in the normal sense of the art, e.g. an RNA polymerase binding site.

Enhanced expression of the nucleotide sequence encoding a flagellin of the present description may also be achieved by the selection of heterologous regulatory regions, e.g. promoter, secretion leader and terminator regions.

In one embodiment, the nucleotide sequence according to the present description is operably linked to at least a promoter.

Other promoters may even be used to direct expression of the polypeptide of the present description.

Examples of suitable promoters for directing the transcription of the nucleotide sequence in a bacterial, fungal or yeast host are well known in the art.

The promoter can additionally include features to ensure or to increase expression in a suitable host. For example, the features can be conserved regions such as a Pribnow Box or a TATA box.

Constructs

The term "construct" - which is synonymous with terms such as "conjugate", "cassette" and "hybrid" - includes a nucleotide sequence for use according to the present invention directly or indirectly attached to a promoter.

An example of an indirect attachment is the provision of a suitable spacer group such as an intron sequence, such as the Sh1-intron or the ADH intron, intermediate the promoter and the nucleotide sequence of the present description. The same is true for the term "fused" in relation to the present description which includes direct or indirect attachment. In some cases, the terms do not cover the natural combination of the nucleotide sequence coding for the protein ordinarily associated with the wild type gene promoter and when they are both in their natural environment.

The construct may even contain or express a marker, which allows for the selection of the genetic construct.

For some applications, the construct of the present description comprises at least the nucleotide sequence of the present description operably linked to a promoter.

Host cells

The term "host cell" - in relation to the present description includes any cell that comprises the nucleotide sequence and/or an expression vector as described herein. Typically the host cell is capable of the recombinant production of a protein having the specific properties as defined herein.

Examples of host cells include bacteria such as *Roseburia* spp. and competent cells. Examples of *Roseburia* spp are *Roseburia hominis*, *Roseburia cecicola*, *Roseburia faecis*, *Roseburia intestinalis*, and *Roseburia inulinivorans*. Examples of competent cells include competent *E. coli* cells (such as *E. coli* BL21(DE3) pLysS and/or *E. coli* B21 Rosetta).

Thus, a further embodiment of the present description provides host cells transformed or transfected with a nucleotide sequence of the present description. In addition, or alternatively, a further embodiment of the present description provides host cells transformed or transfected with a nucleotide sequence (e.g. a promoter such as a heterologous promoter or an exogenous) that is capable of upregulating (overexpressing) the expression of a nucleotide sequence (e.g. gene - such as a homologous gene or an endogenous gene) encoding a flagellin of the present description when compared to the equivalent microorganism prior to transformation. The cells will be chosen to be compatible with the said vector and may for example be bacterial (e.g. prokaryotic), fungal, or yeast cells.

The nucleotide sequence encoding the flagellin of the present description may be heterologous or homologous to the host cell. Typically when the nucleotide sequence encoding the flagellin is homologous to the host cell the host cell comprises multiple copies of the nucleotide sequence. In addition or alternatively, the nucleotide sequence encoding the flagellin is operably linked to a heterologous promoter; typically said promoter is capable of upregulating (overexpressing) the homologous nucleotide sequence encoding the flagellin.

In one example, the host cell comprises the nucleotide sequence coding for the flagellin polypeptide of the present description (such as a homologous or endogenous nucleotide

sequence) under the control of a heterologous or exogenous promoter.

In one embodiment, the host cell is transformed or transfected with one or more nucleotide sequences that encodes at least one flagellin (e.g. *Roseburia* flagellin) selected from the group consisting of Fla1, Fla2, Fla3 and Fla4. In another embodiment, the host cell is transformed or transfected with one or more nucleotide sequences that encode at least one flagellin (e.g. *Roseburia* flagellin) selected from the group consisting of Fla2, Fla1 and Fla4. In a further embodiment, the host cell is transformed or transfected with a nucleotide sequences that encodes the flagellin (e.g. *Roseburia* flagellin) Fla2.

The host cell may comprise multiple copies of polynucleotide sequences encoding *Roseburia* flagellins.

A host cell may comprise at least 1, 2, 3, 4 or 5 polynucleotide sequences encoding at least 1, 2, 3, 4 or 5 *Roseburia* flagellins.

In one embodiment, the host cell comprises a polynucleotide sequence encoding at least one *Roseburia* flagellin derived or derivable from one *Roseburia* species and at least one further polynucleotide sequence encoding at least one *Roseburia* flagellin derived or derivable from a different *Roseburia* species. For example, the host cell comprises at least one polynucleotide sequence encoding an *R. hominis* flagellin (e.g. Fla1 or Fla2) and at least one polynucleotide sequence encoding a *R. intestinalis* flagellin (e.g. Fla1 or Fla 2).

A host cell may comprise at least 1, 2, 3, 4 or 5 expression vectors comprising at least 1, 2, 3, 4 or 5 polynucleotide sequences encoding *Roseburia* flagellin.

In one embodiment, the host cell comprises an expression vector comprising a polynucleotide sequence encoding *Roseburia* flagellins derived or derivable from one *Roseburia* species and at least one further expression vector comprising a polynucleotide sequence encoding *Roseburia* flagellins derived or derivable from a different *Roseburia* species. For example, the host cell comprises at least one expression vector comprising a polynucleotide sequence encoding an *R. hominis* flagellin (e.g. Fla1 or Fla2) and at least one expression vector comprising a polynucleotide sequence encoding a *R. intestinalis* flagellin (e.g. Fla1 or Fla 2).

The nucleotide sequence encoding the flagellin may be endogenous or exogenous to the host cell. Typically when the nucleotide sequence encoding the flagellin is endogenous to

the host cell the host cell comprises multiple copies of the nucleotide sequence. In addition or alternatively, the nucleotide sequence encoding the flagellin is operably linked to an exogenous promoter; typically said promoter is capable of upregulating (overexpressing) the endogenous nucleotide sequence encoding the flagellin.

Examples of suitable bacterial host organisms are gram positive or gram negative bacterial species.

In one embodiment, the host cell is a microorganism.

In one embodiment, the host cell is a lactic acid bacterium species, *Lactococcus* species, a *Bifidobacterium* species, a *Lactobacillus* species or a *Propionibacterium* species.

Examples of lactic acid bacteria include, but are not limited to, *Lactobacillus* spp., *Leuconostoc* spp., *Pediococcus* spp., *Lactococcus* spp., *Streptococcus* spp., *Aerococcus* spp., *Carnobacterium* spp., *Enterococcus* spp., *Oenococcus* spp., *Sporolactobacillus* spp., *Tetragenococcus* spp., *Vagococcus* spp., and *Weisella* spp.

Examples of *Lactobacillus* spp include *Lactobacillus paracasei*, *L. acidophilus*, *L. fermentum*, *L. brevis*, *L. gasseri*, *L. plantarum*, *L. bulgaricus*, *L. helveticus*, *L. reuteri*, *L. casei*, *L. jensenii*, *L. rhamnosus*, *L. crispatus*, *L. johnsonii*, *L. salivarius*, *L. acetotolerans*, *L. acidifarinae*, *L. acidipiscis*, *L. agilis*, *L. algidus*, *L. alimentarius*, *L. amylolyticus*, *L. amylophilus*, *L. amylotrophicus*, *L. amylovorus*, *L. animalis*, *L. antri*, *L. apodemi*, *L. aviarius*, *L. bifermentans*, *L. buchneri*, *L. camelliae*, *L. cateniformis*, *L. ceti*, *L. coleohominis*, *L. collinoides*, *L. composti*, *L. concavus*, *L. coryniformis*, *L. crustorum*, *L. curvatus*, *L. delbrueckii subsp. delbrueckii*, *L. delbrueckii subsp. bulgaricus*, *L. delbrueckii subsp. lactis*, *L. dextrinicus*, *L. diolivorans*, *L. equi*, *L. equigenerosi*, *L. farraginis*, *L. farciminis*, *L. fornicalis*, *L. fructivorans*, *L. frumenti*, *L. fuchuensis*, *L. galUnarum*, *L. gastricus*, *L. ghanensis*, *L. graminis*, *L. hammesii*, *L. hamsteri*, *L. harbinensis*, *L. hayakitensis*, *L. hilgardii*, *L. homohiochii*, *L. iners*, *L. ingluviei*, *L. intestinalis*, *L. kalixensis*, *L. kefiranoferens*, *L. kefir*, *L. kimchii*, *L. kitasatonis*, *L. kunkei*, *L. leichmannii*, *L. lindneri*, *L. malefermentans*, *L. mail*, *L. manihotivorans*, *L. mindensis*, *L. mucosae*, *L. murinus*, *L. nagelii*, *L. namurensis*, *L. nantensis*, *L. oligofermentans*, *L. oris*, *L. panis*, *L. pantheris*, *L. parabrevis*, *L. parabuchneri*, *L. paracollinoides*, *L. parafarraginis*, *L. parakefiri*, *L. paralimentarius*, *L. paraplanarum*, *L. pentosus*, *L. perolens*, *L. pontis*, *L. psittaci*, *L. rennini*, *L. rimae*, *L. rogosae*, *L. rossiae*, *L. ruminis*, *L. saerimneri*, *L. sakei*, *L. sanfranciscensis*, *L. satsumensis*, *L. secaliphilus*, *L. sharpeae*, *L. siliginis*, *L. spicheri*, *L. suebicus*, *L. thailandensis*, *L. ultunensis*, *L. vaccinostercus*, *L. vaginalis*, *L. versmoldensis*, *L. vini*, *L.*

vitulinus, *L. zeae*, and *L. zymae*.

Examples of *Propionibacterium* include, but are not limited to *Propionibacterium freudenreichii* subsp. *shermanii* (PAB), *Propionibacterium acidifaciens*, *Propionibacterium acidipropionici*, *Propionibacterium acnes*, *Propionibacterium australiense*, *Propionibacterium avidum*, *Propionibacterium cyclohexanicum*, *Propionibacterium freudenreichii* subsp. *freudenreichii*, *Propionibacterium granulosum*, *Propionibacterium jensenii*, *Propionibacterium microaerophilum*, *Propionibacterium propionicum*, and *Propionibacterium thoenii*.

In one embodiment, the *Propionibacterium* is *Propionibacterium freudenreichii* subsp. *shermanii* (PAB).

Examples of *Bifidobacterium* include, but are not limited to, *Bifidobacterium adolescentis*, *B. breve*, *B. longum*, *B. animalis*, *B. infantis*, *B. thermophilum*, *B. bifidum*, *Bifidobacterium catenulatum*, *Bifidobacterium pseudocatenulatum*, *Bifidobacterium angulatum* and *B. lactis*.

In another embodiment, the host cell is a Firmicute – for example a *Roseburia* species such as *Roseburia hominis*, *Roseburia cecicola*, *Roseburia faecis*, *Roseburia intestinalis*, or *Roseburia inulinivorans*). In one embodiment, the host cell comprises at least one heterologous polynucleotide sequence encoding a *Roseburia* flagellin. In addition or alternatively, the host cell comprises at least two copies of a homologous polynucleotide sequence encoding a *Roseburia* flagellin; for instance, the host cell comprises at least 3, 4 or 5 homologous copies of a polynucleotide sequence encoding a *Roseburia* flagellin.

Depending on the nature of the nucleotide sequence encoding the polypeptide of the present description, and/or the desirability for further processing of the expressed protein, eukaryotic hosts such as yeasts or other fungi may be used. In general, yeast cells are used over fungal cells because they are easier to manipulate. However, some proteins are either poorly secreted from the yeast cell, or in some cases are not processed properly (e.g. hyperglycosylation in yeast). In these instances, a different fungal host organism should be selected.

The use of suitable host cells - such as yeast, fungal and plant host cells - may provide for post-translational modifications (e.g. myristoylation, glycosylation, truncation, lipidation and tyrosine, serine or threonine phosphorylation) as may be needed to confer optimal biological activity on recombinant expression products of the present description.

The host cell may be a protease deficient or protease minus strain. This may for example be the protease deficient strain *Aspergillus oryzae* JaL 125 having the alkaline protease gene named "alp" deleted. This strain is described in WO97/35956.

The term "host cell" does not cover native nucleotide coding sequences in their natural environment when they are under the control of their native promoter which is also in its natural environment.

Organism

The term "organism" in relation to the present description includes any organism that could comprise the nucleotide sequence coding for the polypeptide according to the present description and/or products obtained therefrom, and/or wherein a promoter can allow expression of the nucleotide sequence according to the present description when present in the organism.

Suitable organisms may include a bacterium (such as a prokaryote), a fungus, a yeast or a plant.

The term "transgenic organism" in relation to the present description includes any organism that comprises the nucleotide sequence coding for the polypeptide according to the present description and/or the products obtained therefrom, and/or wherein a promoter can allow expression of the nucleotide sequence according to the present description within the organism. In one embodiment the nucleotide sequence is incorporated in the genome of the organism.

The term "transgenic organism" does not cover native nucleotide coding sequences in their natural environment when they are under the control of their native promoter which is also in its natural environment.

Therefore, the transgenic organism of the present description includes an organism comprising any one of, or combinations of, the nucleotide sequence coding for the polypeptide according to the present description, constructs according to the present description, vectors according to the present description, plasmids according to the present description, cells according to the present description, or the products thereof.

For example the transgenic organism may comprise the nucleotide sequence coding for the

polypeptide of the present description (such as a homologous nucleotide sequence) under the control of a heterologous promoter.

Transformation of host cells/organism

As indicated earlier, the host organism can be a prokaryotic or a eukaryotic organism. Examples of suitable prokaryotic hosts include *Roseburia hominis*, *Roseburia cecicola*, *Roseburia faecis*, *Roseburia intestinalis*, *Roseburia inulinivorans*, *E. coli* and *Bacillus subtilis*.

Teachings on the transformation of prokaryotic hosts is well documented in the art, for example see Sambrook *et al* (Molecular Cloning: A Laboratory Manual, 2nd edition, 1989, Cold Spring Harbor Laboratory Press). If a prokaryotic host is used then the nucleotide sequence may need to be suitably modified before transformation - such as by removal of introns.

Filamentous fungi cells may be transformed using various methods known in the art – such as a process involving protoplast formation and transformation of the protoplasts followed by regeneration of the cell wall in a manner known. The use of *Aspergillus* as a host microorganism is described in EP 0 238 023.

Another host organism can be a plant. A review of the general techniques used for transforming plants may be found in articles by Potrykus (*Annu Rev Plant Physiol Plant Mol Biol* [1991] 42:205-225) and Christou (Agro-Food-Industry Hi-Tech March/April 1994 17-27). Further teachings on plant transformation may be found in EP-A-0449375.

General teachings on the transformation of fungi, yeasts and plants are presented in following sections.

Transformed fungus

A host organism may be a fungus - such as a mould. Examples of suitable such hosts include any member belonging to the genera *Thermomyces*, *Acremonium*, *Aspergillus*, *Penicillium*, *Mucor*, *Neurospora*, *Trichoderma* and the like.

In one embodiment, the host organism may be a filamentous fungus.

Transforming filamentous fungi is discussed in US-A-5741665 which states that standard

techniques for transformation of filamentous fungi and culturing the fungi are well known in the art. An extensive review of techniques as applied to *N. crassa* is found, for example in Davis and de Serres, *Methods Enzymol* (1971) **17A**: 79-143.

5 Further teachings which may also be utilised in transforming filamentous fungi are reviewed in US-A-5674707.

10 In addition, gene expression in filamentous fungi is taught in in Punt *et al.* (2002) Trends Biotechnol 2002 May;20(5):200-6, Archer & Peberdy Crit Rev Biotechnol (1997) 17(4):273-306.

The present description encompasses the production of transgenic filamentous fungi according to the present description prepared by use of these standard techniques.

15 In one aspect, the host organism can be of the genus *Aspergillus*, such as *Aspergillus niger*.

A transgenic *Aspergillus* according to the present description can also be prepared by following, for example, the teachings of Turner G. 1994 (Vectors for genetic manipulation. In: Martinelli S.D., Kinghorn J.R.(Editors) *Aspergillus*: 50 years on. Progress in industrial microbiology vol 29. Elsevier Amsterdam 1994. pp. 641-666).

Transformed yeast

In another embodiment, the transgenic organism can be a yeast.

25 A review of the principles of heterologous gene expression in yeast are provided in, for example, *Methods Mol Biol* (1995), 49:341-54, and *Curr Opin Biotechnol* (1997) Oct;8(5):554-60

30 In this regard, yeast – such as the species *Saccharomyces cerevisi* or *Pichia pastoris* (see FEMS Microbiol Rev (2000 24(1):45-66), may be used as a vehicle for heterologous gene expression.

35 A review of the principles of heterologous gene expression in *Saccharomyces cerevisiae* and secretion of gene products is given by E Hinchcliffe E Kenny (1993, "Yeast as a vehicle for the expression of heterologous genes", *Yeasts*, Vol 5, Anthony H Rose and J Stuart Harrison, eds, 2nd edition, Academic Press Ltd.).

For the transformation of yeast, several transformation protocols have been developed. For example, a transgenic *Saccharomyces* according to the present description can be prepared by following the teachings of Hinnen *et al.*, (1978, *Proceedings of the National Academy of Sciences of the USA* **75**, 1929); Beggs, J D (1978, *Nature*, London, **275**, 104); and Ito, H *et al* (1983, *J Bacteriology* 153, 163-168).

The transformed yeast cells may be selected using various selective markers – such as auxotrophic markers dominant antibiotic resistance markers.

Transformed plants/plant cells

A host organism suitable for the present description may be a plant. In this respect, the basic principle in the construction of genetically modified plants is to insert genetic information in the plant genome so as to obtain a stable maintenance of the inserted genetic material. A review of the general techniques may be found in articles by Potrykus (*Annu Rev Plant Physiol Plant Mol Biol* [1991] **42**:205-225) and Christou (Agro-Food-Industry Hi-Tech March/April 1994 17-27).

Direct infection of plant tissues by *Agrobacterium* is a simple technique which has been widely employed and which is described in Butcher D.N. *et al.*, (1980), *Tissue Culture Methods for Plant Pathologists*, eds.: D.S. Ingrams and J.P. Helgeson, 203-208.

Other techniques for transforming plants include ballistic transformation, the silicon whisker carbide technique (see Frame BR, Drayton PR, Bagnaall SV, Lewnau CJ, Bullock WP, Wilson HM, Dunwell JM, Thompson JA & Wang K (1994) Production of fertile transgenic maize plants by silicon carbide whisker-mediated transformation, *The Plant Journal* **6**: 941-948) and viral transformation techniques (e.g. see Meyer P, Heidmann I & Niedenhof I (1992) The use of cassava mosaic virus as a vector system for plants, *Gene* **110**: 213-217).

Further teachings on plant transformation may be found in EP-A-0449375.

Plant cells may be grown and maintained in accordance with well-known tissue culturing methods such as by culturing the cells in a suitable culture medium supplied with the necessary growth factors such as amino acids, plant hormones, vitamins, etc.

In a further aspect, the present description relates to a vector system which carries a nucleotide sequence or construct according to the present description and which is capable of introducing

the nucleotide sequence or construct into the genome of an organism, such as a plant. The vector system may comprise one vector, but it may comprise two vectors. In the case of two vectors, the vector system is normally referred to as a binary vector system. Binary vector systems are described in further detail in Gynheung An *et al.*, (1980), Binary Vectors, *Plant Molecular Biology Manual A3*, 1-19.

One extensively employed system for transformation of plant cells uses the Ti plasmid from *Agrobacterium tumefaciens* or a Ri plasmid from *Agrobacterium rhizogenes* An *et al.*, (1986), *Plant Physiol.* 81, 301-305 and Butcher D.N. *et al.*, (1980), *Tissue Culture Methods for Plant Pathologists*, eds.: D.S. Ingrams and J.P. Helgeson, 203-208. After each introduction method of the desired promoter or construct or nucleotide sequence according to the present description in the plants, the presence and/or insertion of further DNA sequences may be necessary. If, for example, for the transformation the Ti- or Ri-plasmid of the plant cells is used, at least the right boundary and often however the right and the left boundary of the Ti- and Ri-plasmid T-DNA, as flanking areas of the introduced genes, can be connected. The use of T-DNA for the transformation of plant cells has been intensively studied and is described in EP-A-120516; Hoekema, in: The Binary Plant Vector System Offset-drukkerij Kanters B.B., Alblaserdam, 1985, Chapter V; Fraley, *et al.*, *Crit. Rev. Plant Sci.*, 4:1-46; and An *et al.*, *EMBO J.* (1985) 4:277-284.

Culturing and production

Host cells transformed with the nucleotide sequence of the present description and/or an expression vector of the present description may be cultured under conditions conducive to the production of the encoded polypeptide and which facilitate recovery of the polypeptide from the cells and/or culture medium.

The medium used to cultivate the cells may be any conventional medium suitable for growing the host cell in questions and obtaining expression of the polypeptide.

The protein produced by a recombinant cell may be displayed on the surface of the cell.

The protein may be secreted from the host cells and may conveniently be recovered from the culture medium using well-known procedures.

Secretion

In some embodiments, the protein is secreted from the expression host into the culture medium from where the protein may be recovered. According to the present description, the secretion leader sequence may be selected on the basis of the desired expression host. Hybrid signal sequences may also be used with the context of the present description.

Typical examples of heterologous secretion leader sequences are those originating from the fungal amyloglucosidase (AG) gene (*glaA* - both 18 and 24 amino acid versions e.g. from *Aspergillus*), the α -factor gene (yeasts e.g. *Saccharomyces*, *Kluyveromyces* and *Hansenula*) or the α -amylase gene (*Bacillus*).

By way of example, the secretion of heterologous proteins in *E. coli* is reviewed in Methods Enzymol (1990) 182:132-43.

Detection

A variety of protocols for detecting and measuring the expression of the amino acid sequence are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and fluorescent activated cell sorting (FACS).

A wide variety of labels and conjugation techniques are known by those skilled in the art and can be used in various nucleic and amino acid assays.

A number of companies such as Pharmacia Biotech (Piscataway, NJ), Promega (Madison, WI), and US Biochemical Corp (Cleveland, OH) supply commercial kits and protocols for these procedures.

Suitable reporter molecules or labels include those radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles and the like. Patents teaching the use of such labels include US-A-3,817,837; US-A-3,850,752; US-A-3,939,350; US-A-3,996,345; US-A-4,277,437; US-A-4,275,149 and US-A-4,366,241.

Also, recombinant immunoglobulins may be produced as shown in US-A-4,816,567.

Fusion proteins

The amino acid sequence for use according to the present description may be produced as a fusion protein, for example to aid in extraction and purification. Examples of fusion protein partners include glutathione-S-transferase (GST), 6xHis, GAL4 (DNA binding and/or transcriptional activation domains) and (β -galactosidase). It may also be convenient to include a proteolytic cleavage site between the fusion protein partner and the protein sequence of interest to allow removal of fusion protein sequences.

Typically, the fusion protein will not hinder the activity of the protein sequence.

Gene fusion expression systems in *E. coli* have been reviewed in *Curr Opin Biotechnol* (1995) 6(5):501-6.

In another embodiment, the amino acid sequence may be ligated to a heterologous sequence to encode a fusion protein. For example, for screening of peptide libraries for agents capable of affecting the substance activity, it may be useful to encode a chimeric substance expressing a heterologous epitope that is recognised by a commercially available antibody.

The present invention is further described by way of the following non-limiting examples.

EXAMPLES

The practice of the present invention will employ, unless otherwise indicated, conventional techniques of chemistry, molecular biology, microbiology, recombinant DNA and immunology, which are within the capabilities of a person of ordinary skill in the art. Such techniques are explained in the literature. See, for example, J. Sambrook, E. F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Second Edition, Books 1-3, Cold Spring Harbor Laboratory Press; Ausubel, F. M. *et al.* (1995 and periodic supplements; *Current Protocols in Molecular Biology*, ch. 9, 13, and 16, John Wiley & Sons, New York, N.Y.); B. Roe, J. Crabtree, and A. Kahn, 1996, *DNA Isolation and Sequencing: Essential Techniques*, John Wiley & Sons; J. M. Polak and James O'D. McGee, 1990, *In Situ Hybridization: Principles and Practice*; Oxford University Press; M. J. Gait (Editor), 1984, *Oligonucleotide Synthesis: A Practical Approach*, Irl Press; D. M. J. Lilley and J. E. Dahlberg, 1992, *Methods of Enzymology: DNA Structure Part A: Synthesis and Physical Analysis of*

DNA Methods in Enzymology, Academic Press; and E. M. Shevach and W. Strober, 1992 and periodic supplements, *Current Protocols in Immunology*, John Wiley & Sons, New York, NY. Each of these general texts is herein incorporated by reference.

Material and Methods

ROSEBURIA GENUS FLAGELLINS

Flagellin diversity

Flagellins from *Roseburia* genus bacteria, in particular, *Roseburia hominis* and *Roseburia intestinalis* were cloned, expressed, purified and analysed.

Figure B1A and B1B show an SDS-Analysis of recombinant flagellins.

Flagellin protein nomenclature:

***Roseburia* Genus**

***Roseburia* Species**

Roseburia hominis

Roseburia hominis FlaA1 (This is also referred herein as RhFlaA1 or Rh1)

Roseburia hominis FlaA2 (This is also referred herein as RhFlaA2 or Rh2)

Roseburia intestinalis

Roseburia intestinalis FlaA1 (This is also referred herein as RiFlaA1 or Ri1 or RI1)

Roseburia intestinalis FlaA2 (This is also referred herein as RiFlaA2 or Ri2 or RI2)

Roseburia intestinalis FlaA3 (This is also referred herein as RiFlaA3 or Ri3 or RI3)

Roseburia intestinalis FlaA4 (This is also referred herein as RiFlaA1 or Ri4 or RI4)

See:

ELY, B., ELY, T.W., CRYMES, W.B.,JR and MINNICH, S.A., 2000. A family of six flagellin genes contributes to the *Caulobacter crescentus* flagellar filament. *Journal of Bacteriology*, **182**(17), pp. 5001-5004.

IBRAHIM, G.F., FLEET, G.H., LYONS, M.J. and WALKER, R.A., 1985. Method for the isolation of highly purified *Salmonella* flagellins. *Journal of clinical microbiology*, **22**(6), pp.

1040-1044.

NEVILLE, B.A., FORDE, B.M., CLAEISSON, M.J., DARBY, T., COGHLAN, A., NALLY, K., ROSS, R.P. and O'TOOLE, P.W., 2012. Characterization of pro-inflammatory flagellin proteins produced by *Lactobacillus ruminis* and related motile *Lactobacilli*. *PloS one*, **7**(7), pp. e40592.

NG, S.Y., CHABAN, B. and JARRELL, K.F., 2006. Archaeal flagella, bacterial flagella and type IV pili: a comparison of genes and posttranslational modifications. *Journal of Molecular Microbiology and Biotechnology*, **11**(3-5), pp. 167-191.

WATSON, R.O. and GALAN, J.E., 2005. Signal transduction in *Campylobacter jejuni*-induced cytokine production. *Cellular microbiology*, **7**(5), pp. 655-665.

Bacterial growth conditions

R. hominis A2-183^T (=DSM 16839^T=NCIMB 14029^T) was grown anaerobically at 37°C in YCFA media. The culture was spun down and the pellet was resuspended in one mL of YCFA media, supplemented with 2% cysteine (w/v, Sigma-Aldrich) and 3% ascorbic acid (w/v, Sigma-Aldrich).

R. intestinalis L1-82^T (=DSM 14610^T=NCIMB 13810^T) was grown anaerobically at 37°C in YCFA media. The culture was spun down and the pellet was resuspended in one mL of YCFA media, supplemented with 2% cysteine (w/v, Sigma-Aldrich) and 3% ascorbic acid (w/v, Sigma-Aldrich).

Mice

C3H/HeN and C57Bl/6 were purchased from Harlan Laboratories. GF C3H/HeN were provided and maintained in the INRA gnotobiotic rodent breeding facility at Jouy-en-Josas (ANAXEM platform, Institut Micalis, INRA, Jouy-en-Josas, France). GF TLR5KO and wild type C57Bl/6 were provided by Andrew Gewirtz (Center for Inflammation, Immunity, and Infection and Department of Biology, Georgia State University, Atlanta, GA 30303, USA) and maintained in the INRA gnotobiotic rodent breeding facility at Jouy-en-Josas. Conventional TLR5KO and wild type BOY/J were provided by Adam Cunningham (MRC Centre for Immune Regulation, Institute of Microbiology and Infection, Division of Immunity and Infection University of Birmingham UK) The management and experimental procedures were approved by the respective Local Ethical Review Committees.

Animal experiments

Germfree animal experiments were performed in the INRA gnotobiotic rodent breeding facility at Jouy-en-Josas (ANAXEM platform, Institut Micalis, INRA, Jouy-en-Josas, France).

GF C3H/HeN male mice were allocated into control ($N=8$) and treatment ($N=10$) groups and caged individually in plastic isolators. At day 0, 1 and 2, animals in the treatment group were given 100 μ L of *R. hominis* culture by gavage, while control animals were given 100 μ L YCFA media. Ileum, ascending colon and caecum samples were collected at 14d and 28d.

Six GF C3H/HeN male mice were treated with *E. coli* MG1655 (K12) as described above, and three animals were sacrificed at 10d and 22d to give $N=3$. Three GF TLR5KO mice and three C57Bl/6 WT mice were inoculated with *R. hominis* culture as described above to evaluate the functional importance of *R. hominis* flagellins. After 28d these animals were sacrificed together with their GF counterparts. Twenty-two female C57BL/6 mice were dosed daily with 50 μ L of 10^9 CFU *R. hominis* for 14 days. Control animals were dosed with culture medium alone. From day 8, mice were given DSS (MW 50kDa, 30g/l) in their drinking water for 6 days. The animals were euthanized on day 14 and tissue sampling was performed as described above.

Tissue culture experiments

Caco-2 (*Homo sapiens* epithelial colorectal adenocarcinoma cells) and HT29 (*Homo sapiens* colorectal adenocarcinoma) cells were grown in transwell plates within an anaerobic workstation. *R. hominis* A2-183 culture or *R. intestinalis* L1-82^T was harvested at exponential phase, and 100 μ L of bacterial suspension (10^8 CFU/mL) was added to experimental wells. Bacterial (non-adherent and adherent) and eukaryotic cells (Caco-2 and HT-29) were harvested after 2h and 4h incubation and stored in RNeasy lysis buffer. For tissue culture experiments with recombinant flagellins, 5×10^4 Caco-2 cells were grown in 24-well plates at 37°C in a 75% humidified atmosphere of 5% CO₂. The cells reached confluence on day 5-6 and were used three days post-confluence. The cells were incubated with recombinant flagellins at a final concentration of 100 ng/ μ l for 2hr at 37°C in a 75% humidified atmosphere of 5% CO₂.

FISH analysis

FISH analysis was performed on neutral buffered formalin-fixed gut tissue sections using a general bacterial probe Eub338 and a newly designed *R. hominis* A2-183-specific probe and a *R. intestinalis* L1-82^T specific probe.

R. hominis library construction

R. hominis chromosomal DNA for small-size library construction and pyrosequencing was isolated using an UltraClean™ Microbial DNA Isolation Kit (Mo Bio Laboratories Inc) and high-molecular-weight DNA for fosmid libraries was isolated using a Wizard Genomic DNA Purification kit (Promega).

Microarray analyses

Bacterial microarray

Bacterial RNA was isolated from mouse caecum contents and further processed using commercial kits following the manufacturer's recommendations. PCR products amplified from 6000 clones in the *E. coli* plasmid library of *R. hominis* were arrayed in duplicate on aminosilane-coated microscope slides (Corning) using a MicroGrid II TAS (BioRobotics).

Mouse microarray analysis

Total RNA was extracted from ileum and ascending colon tissue, processed into biotin-labelled cRNA/aRNA (depending on Affymetrix kit used), and hybridized to the GeneChip NuGO Mouse Array and GeneChip Mouse Genome Array (Affymetrix) using standard techniques. Data analysis was performed with the software packages R (<http://www.r-project.org>) and Bioconductor (<http://www.bioconductor.org>).

RT-PCR analysis

The *R. hominis*-specific primers 5'-CCCACTGACAGAGTATGTAATGTAC-3' and 5'-GCACCACCTGTCACCAC-3' were used for PCR analysis of faecal samples to validate gut colonization levels. Real-time PCR analysis was performed using a 7500 Fast Real-Time PCR System (Applied Biosystems) with the Power SYBR Green PCR Master Mix (Applied Biosystems). All samples were run in triplicate. *GyrA* was used as a reference gene for normalization. For host gene expression, total eukaryotic RNA isolated from the ileum and ascending colon was reverse-transcribed into cDNA using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Real-time PCR analysis was performed using a 7500 Fast Real-Time PCR System (Applied Biosystems) with the QuantiFast SYBR Green PCR Kit (Qiagen) and QuantiTect Primer Assays (Qiagen). All samples were run in triplicate. *Hprt* was selected as a reference gene for normalization. All RT-PCR data were analyzed on a logarithmic scale with base 2 by one-way ANOVA with a significance cut-off of $P < 0.05$. Differences were back-transformed to calculate fold changes.

Western blot

Immuno-purified rabbit polyclonal antibodies against *Roseburia hominis* FlaA1 and FlaA2 was produced as described in Duck *et al.* (Duck *et al.* 2007). For the western blot, *R. hominis* was grown in the presence of varying amounts (0.01g to 1g diet/10 mL of culture) of UV irradiated standard mouse chow for 3hrs, filtered to remove dietary components and diluted in Laemmli buffer containing 8M urea. Samples were loaded on a NuPAGE® Novex® 4-12% Bis-Tris gel (Invitrogen) and electrophoresed, followed by further processing using the WesternBreeze Chromogenic Immunodetection System (Invitrogen). FlaA1 and FlaA2

antibodies were used at 1:1000 and loading control anti-DNA gyrase A (Abcam) at 1:300, followed by alkaline phosphatase conjugated anti-rabbit. Detection was by substrate colour development relative to loading control colour development.

Immunofluorescence

Immuno-localization of *R. hominis* flagellin was examined in colon contents of mice using specific antisera raised against defined peptide sequences from both FlaA1 and FlaA2 flagellin proteins. Gut content smears were fixed in pre-cooled methanol, incubated with anti-FlaA1 or anti-FlaA2 rabbit antisera (CovaLabs) at 4°C overnight and visualized using Alexa donkey anti rabbit 488 (Molecular Probes).

T cell markers were examined on sequential 8 µm cryosection. Fixed tissue sections were incubated with Ly6G-FITC, CD3-FITC, CD11b-FITC (BD Biosciences), double-labelled with FoxP3 (Abcam) and CD3-FITC (BD Biosciences) primary antibody or isospecific IgG. Sections were counterstained with DAPI and mounted with Vectashield (Vector Laboratories). For quantification of positive cells, a minimum of five fields of view from each mouse section was examined.

Cloning and purification of recombinant flagellins

Flagellin genes were isolated from liquid bacterial cultures of *R. hominis*, *R. intestinalis*, *S. typhimurium*, *S. enteritidis*, *Eubacterium rectale* 33656 and *E. coli* K12 by PCR amplification and purification. Caco-2 cells were incubated with recombinant flagellins at a final concentration of 100 ng/µL for 2h at 37°C in a 75% humidified atmosphere of 5% CO₂.

Isolation of intestinal and MLN cells

Cells were isolated from the small intestine and mesenteric lymph node as previously described with minor modifications (Monteleone *et al.* 2008). Briefly, cellular suspensions were incubated with 100U/mL collagenase VIII (Sigma-Aldrich) in RPMI supplemented with 20% FBS at 37°C for 20 min (mesenteric lymph nodes) or 1 hour (intestinal pieces). Single cell suspensions were then analyzed by flow cytometry (as described).

Generation of bone marrow-derived dendritic cells and cultures

Bone marrow was harvested from femur and tibia of C3H/HeN and C57Bl6 mice. For GM-CSF-derived dendritic cells, bone marrow cells were resuspended at 1x10⁶/mL in RPMI supplemented with 10% FCS and 20ng/mL rmGM-CSF and seeded at 10mL/plate in 100mm² tissue culture plates. After three days culture, loosely adherent cells were collected and replated with GM-CSF supplemented media at 1x10⁶/mL in 12 well tissue culture plates.

At day 5, cells were stimulated with 100ng/mL flagellins before being harvested on day 6. For Flt3L-derived dendritic cells, bone marrow cells were resuspended at 2×10^6 /mL in RPMI supplemented with 10% FCS and 200ng/mL rmFlt3 and seeded at 2mL/well in 12-well tissue culture plates. Cells were cultured for 10 days with Flt3 supplemented media added to each well on day 4. At day 9, cells were stimulated with 100ng/mL flagellins before being harvested on day 10 and analyzed by flow cytometry.

Flow Cytometry

Single cell suspensions of lamina propria cells and dendritic cells were incubated in blocking buffer (containing serum and CD16/CD32 antibody) at 4°C for 15 min prior to staining with specific fluorochrome conjugated antibodies. Lamina propria cells were labelling with antibodies to mouse CD4-FITC and CD25-APC (eBioscience), CD8-APC-Cy7 and CD3-PerCP (Biolegend), and B220-BV570 (BD Biosciences). Intracellular FoxP3 labelling was performed after extracellular staining and cell fixation/permeabilisation according to the manufacturer's instructions (eBioscience). GMCSF-derived dendritic cells were labelled with antibodies CD11c-PE-Cy7, CD11-PerCP Cy5, I-A/I-E -APC-Cy7, CD80-PE, CD86-APC, CD8-FITC, B220-BV570. Flt3 derived dendritic cells were labelled with CD11c-PE-Cy7-, CD11b- or Siglec-H-PerCP Cy5-, I-A/I-E -APC-Cy7, CD317 -PE, CD40-APC, CD103-FITC, B220-BV570 Cells were analyzed using a FACSArria (BD Biosciences) and FlowJo software version 7.2.5.

Cytometric Bead Array (CBA)

Bone marrow cells were isolated from femur and tibia with RPMI media of C3H/HeN and C57Bl/6 mice and Flt3L-expanded as described previously. Cells were stimulated with 100 ng/mL flagellin (e.g. *Roseburia* flagellin) after 9 days of culture, and supernatant was collected on day 10. The experiment was performed on three separate occasions to create N=3.

CBA analysis was performed on cell supernatants using the Cytometric Bead Array Mouse Enhanced Sensitivity Master Buffer Kit (BD Biosciences). Standards and samples were loaded onto a 96-well plate for measurement in a FACSArray (BD Biosciences). Results were analyzed using BD FCAP software (BD Biosciences).

Histology

Ascending colon tissue samples were fixed in neutral buffered formalin (Sigma), embedded in cold-curing resin, and 4µm tissue sections were stained using standard haematoxylin/eosin methods. A complete transverse cross sectional area of colon from each

animal was imaged at x200 magnification on a Zeiss Axioskop microscope using a QImaging camera controlled by Image Pro Plus software. Each field of view was then scored from 0-4 according to a method based on Berg *et al.* (Berg *et al.* 1996). The mean percentage of fields of view at a given grade was calculated and treatment groups were compared using student t-test analysis.

Further detailed protocols are described in the supplementary Materials and Methods.

Supplementary Information (SI) Materials and Methods

Bacterial growth conditions

R. hominis A2-183^T (=DSM 16839^T=NCIMB 14029^T) was grown anaerobically on synthetic YCFA or complex M2GSC media. Culture was inoculated from frozen stock into Hungate tubes and incubated overnight at 37°C. Bacteria were then grown on M2GSC agar plates for 48h in a MACS-MG-1000 anaerobic workstation (Don Whitley Scientific) under 80% N₂, 10% CO₂, and 10% H₂ at 37°C. The effect of mucin was investigated by adding 0.5% (w/v) mucin from porcine stomach type III (Sigma-Aldrich) to the YCFA medium.

For colonization of germfree (GF) mice, *R. hominis* was grown in YCFA media overnight at 37°C. The culture was spun down and the pellet was resuspended in one mL of YCFA media, supplemented with 2% cysteine (w/v, Sigma-Aldrich) and 3% ascorbic acid (w/v, Sigma-Aldrich).

Mice

C3H/HeN and C57Bl/6 were purchased from Harlan Laboratories. Mice were housed within HEPA-filtered flexi-film isolators (Bell Isolation Systems) at the University of Aberdeen. GF C3H/HeN were provided and maintained in the INRA gnotobiotic rodent breeding facility at Jouy-en-Josas (ANAXEM platform, Institut Micalis, INRA, Jouy-en-Josas, France). Germfree TLR5KO and wild type C57Bl/6 were provided by Andrew Gewirtz (Center for Inflammation, Immunity, and Infection and Department of Biology, Georgia State University, Atlanta, GA 30303, USA) and maintained in the INRA gnotobiotic rodent breeding facility at Jouy-en-Josas. The management and experimental procedures were approved by the respective Local Ethical Review Committees.

Mouse experiments

Eighteen GF C3H/HeN male mice were allocated into control (*N*=8) and treatment (*N*=10) groups and caged individually in plastic isolators. The mice were fed *ad libitum* on a sterilized

commercial diet (R03-40; UAR). At day 0, 1 and 2, animals in the treatment group were given 100 μ L of *R. hominis* culture by gavage, while control animals were given 100 μ L YCFA media. At 14d and 28d, four control animals and five *R. hominis*-treated animals were sacrificed. The ileum and ascending colon were divided into four equal parts and transferred to RNAlater (Ambion), neutral buffered formalin (NBF; Sigma-Aldrich) or liquid nitrogen. The whole caecum was transferred to RNAlater.

To demonstrate the specificity of the response to *R. hominis*, six GF C3H/HeN male mice were treated with *E. coli* MG1655 (K12), and three animals were sacrificed at 10d and 22d as described above to give N=3.

Three GF TLR5KO mice and three C57Bl/6 WT mice were inoculated with *R. hominis* culture as described above to evaluate the functional importance of *R. hominis* flagellins. After 28d these animals were sacrificed together with their GF counterparts.

Twenty-two female C57BL/6 mice (6 weeks old) were used to evaluate the therapeutic effect of *R. hominis* during DSS-induced colitis. After an acclimatization period of 7-10 days, the mice were dosed daily with 50 μ L of 10^9 CFU *R. hominis* for 14 days. Control animals were dosed with culture medium alone. From day 8, mice were given DSS (MW 50kDa, 30g/l) in their drinking water for 6 days. The animals were euthanized on day 14 and tissue sampling was performed as described above.

Tissue culture experiments

All cell culture reagents, unless specified otherwise, were supplied by Sigma-Aldrich. For tissue culture experiments in anaerobic conditions, 2×10^5 Caco-2 or HT29 cells in 1.5 mL DMEM (high glucose, HEPES) medium supplemented with heat-inactivated fetal bovine serum (Gibco), penicillin, streptomycin, amphotericin B and L-glutamine were seeded into the upper compartments of a six-well transwell plate (Corning). The lower compartments contained 3.0 mL of the same medium. Cells were incubated at 37°C in a 5% CO₂ atmosphere until three days post-confluence, washed with Hanks' solution to remove antibiotics and FCS and stepped down in DMEM supplemented with L-glutamine, sodium selenite and transferrin for 24h without antibiotics. Transwell inserts were then transferred to an anaerobic culture box within the anaerobic workstation at 37°C. The upper compartment of each insert was filled with anaerobic DMEM cell medium, while the lower compartment was filled with oxygenated DMEM.

R. hominis A2-183 culture was harvested at exponential phase by centrifugation at 3,500xg

for 5 min. The pellet was washed and resuspended in 0.8 mL anaerobic DMEM. One hundred microliters of bacterial suspension (10^8 CFU/mL) was added to experimental wells. The control wells received the same amount of medium without bacterial cells. Additional controls included bacterial cells incubated without Caco-2 or HT29 cells. Bacterial and eukaryotic cells were harvested after 2h and 4h incubation. Both non-adherent and adherent bacteria were aspirated and stored in RNAlater (Ambion). The viability of *R. hominis* cells was tested by plating onto YCFA plates. Caco-2 cells or HT-29 cells were harvested from the wells and also stored in RNAlater.

R. intestinalis L1-82^T (=DSM 14610^T=NCIMB 13810^T) culture was harvested at exponential phase by centrifugation at 3,500 $\times$ g for 5 min. The pellet was washed and resuspended in 0.8 mL anaerobic DMEM. One hundred microliters of bacterial suspension (10^8 CFU/mL) was added to experimental wells. The control wells received the same amount of medium without bacterial cells. Additional controls included bacterial cells incubated without Caco-2 or HT29 cells. Bacterial and eukaryotic cells were harvested after 2h and 4h incubation. Both non-adherent and adherent bacteria were aspirated and stored in RNAlater (Ambion). The viability of *R. intestinalis* cells was tested by plating onto YCFA plates. Caco-2 cells or HT-29 cells were harvested from the wells and also stored in RNAlater.

For tissue culture experiments with recombinant flagellins, 5×10^4 Caco-2 cells were seeded in 24-well plates in DMEM (high glucose, HEPES) medium supplemented with heat-inactivated fetal bovine serum (Gibco), penicillin, streptomycin, amphotericin B and L-glutamine at 37°C in a 75% humidified atmosphere of 5% CO₂. The cells reached confluence on day 5-6 and were used 3 days post-confluency. Prior any treatment, cells were washed twice with Hanks' Balanced Salt Solution and kept in DMEM supplemented with L-Glutamine, selenium and transferrin for 24 hours.

***R. hominis* library construction**

R. hominis chromosomal DNA for small-size library construction and pyrosequencing was isolated using an UltraClean™ Microbial DNA Isolation Kit (Mo Bio Laboratories Inc) and high-molecular-weight DNA for fosmid libraries was isolated using a Wizard Genomic DNA Purification Kit (Promega). DNA integrity was checked by gel electrophoresis.

DNA was mechanically sheared using a Nebulizer Kit (Invitrogen) and fractionated by gel electrophoresis. DNA fragments of desired size were excised from the gel and purified using a Wizard® SV Gel and PCR Clean-Up System (Promega). End-repair was done with a DNA Terminator End Repair Kit (Lucigen). 1.5-3.5 kb fragments were cloned using the

CloneSmart® LCamp Kit (Lucigen) and a 4-8 kb library was constructed using the pJAZZ®-OC Vector (Lucigen). Fosmid libraries were constructed using the CopyControl™ Fosmid Library Production Kit (Epicentre Biotechnologies). Colonies were picked using an automated colony picker (BioRobotics BioPick, Genomic Solutions) and archived into 384-well microtitre plates containing 70 µL 2xLB medium supplemented with 10% glycerol and corresponding antibiotic. Cells were grown overnight at 37 °C with shaking and stored at -80 °C.

Sequencing, assembly, and annotation

Templates for sequencing of small-size libraries were generated by PCR using one µL of clone biomass and primers SL1 and SR2 surrounding the cloning site of pSMART-LCAMP. PCR products were purified using Multiscreen PCR Clean-up filter plates (Millipore). Recombinant DNA from the pJAZZ®-OC clones was isolated using the Wizard® SV 96 Plasmid DNA Purification System (Promega). Fosmid DNA was isolated using the FosmidMAX™ DNA Purification Kit (Epicentre). End-reads of DNA fragments from *R. hominis* WGS libraries with different insert sizes were obtained using CEQ8000 (Beckman Coulter) and ABI 3770 (Applied Biosystems) DNA sequencers. Genomic DNA from *R. hominis* was also sequenced using 454 GS20 (454 Life Sciences) and 454 FLX sequencers (Roche). The Sanger and 454 data were assembled with MIRA version 3 (http://chevreux.org/projects_mira.html; (1)). The RAST annotation pipeline (<http://rast.nmpdr.org>; (2)) was used for automatic and manual annotation of the genome and for comparative genomic analyses. The annotated genomic sequence of *R. hominis* A2-183 was submitted to GenBank under the accession number CP003040.

Microarray analyses

Bacterial microarray

Bacterial RNA was isolated from mouse caecum contents using the RNeasy Mini Kit (Qiagen), and further processed with the MICROBEnrich™ Kit (Ambion), the MICROBExpress™ Bacterial mRNA Enrichment Kit (Ambion), and the MessageAmp™ II-Bacteria RNA Amplification Kit (Applied Biosystems). RNA was labelled with either dCTP-Cy3 or dCTP-Cy5 during cDNA synthesis (CyScribe First Strand cDNA Labelling Kit; Amersham). Labeled products were purified using the CyScribe GFX Purification Kit (Amersham). PCR products amplified from 6000 clones in the *E. coli* plasmid RA8 library of *R. hominis* were arrayed in duplicate on aminosilane-coated microscope slides (Corning) using a MicroGrid II TAS (BioRobotics). Amplified fragments of the housekeeping genes *rpoD* and *gyrA* were randomly distributed on the array as controls. Microarray hybridization was performed in the GeneTAC hybridization station (Genomic Solutions). Dye labelling was swapped for a second hybridization, and a separate RNA purification was also labelled and

hybridized twice, to ensure reproducibility and to obtain statistically significant results. In total, four slides were hybridized for each comparison, for a total of 12 hybridizing spots per amplified clone. Fluorescence was measured in two channels using a GeneTAC LS IV (Genomic Solutions) with GeneTac Integrator version 3.0.1 software. Spot intensities were log-transformed and Loess normalization was applied to remove differences in probe labelling and hybridization efficiencies. The one-sample t-test was used on the log-ratio values to test for differential expression. Data was considered significant when fold change > 2 and $P < 0.05$.

Mouse microarray analysis

Ileum and ascending colon tissue was removed from RNAlater and lyzed in Trizol (Invitrogen). RNA was isolated using standard chloroform/isopropanol steps. Total RNA was further purified with the RNeasy Kit (Qiagen), including an RNase-free DNase I (Qiagen) digestion step. RNA integrity was determined using the Agilent 2100 Bioanalyzer (Agilent Technologies). Total RNA was processed into biotin-labelled cRNA using the One-Cycle Target Labeling Kit (Affymetrix) or biotin-labelled aRNA using the 3' IVT Express Kit (Affymetrix). Hybridization to the GeneChip NuGO Mouse Array and GeneChip Mouse Genome Array (Affymetrix) on a GeneChip Fluidics Station 450 (Affymetrix) was performed at the Institute of Medical Sciences Microarray Core Facility (University of Aberdeen, UK). Chips were scanned with an Affymetrix GeneChip Scanner 3000 (Affymetrix). Image quality analysis was performed using Gene Chip Operating Software (GCOS) (Affymetrix). Further data analysis was performed with the freely available software packages R (<http://www.r-project.org>) and Bioconductor (<http://www.bioconductor.org>). The moderated F-test provided by the Bioconductor package limma was used to test for differential expression. Data was considered significant when $P < 0.05$ using the Benjamini and Hochberg false discovery method. Statistical analysis was performed separately for each of the two time-points. All differentially expressed genes ($P < 0.05$) were imported into MetaCore analytical software (GeneGo, St Joseph, MI) to generate pathway maps. Integrated pathway enrichment analysis was performed using the knowledge-based canonical pathways and endogenous metabolic pathways. Ranking of relevant integrated pathways was based on P -values calculated using hypergeometric distribution. P -values represented the probability of a given number of genes from the input list to match a certain number of genes in the map by chance, considering the numbers of genes in the experiment versus the number of genes in the map within the full set of all genes on maps.

Gene Ontology (GO) based functional interpretation of the data was performed using DAVID (<http://david.abcc.ncifcrf.gov>), an expanded version of the original web-accessible program

(3). Significantly different transcripts ($P < 0.05$) were allocated into the GO category 'Biological Process' to unearth patterns of gene expression significantly enriched for specific GO terms.

Microarray data were submitted to the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (accession number GSE25544; <http://www.ncbi.nlm.nih.gov/geo>).

RT-PCR analysis

The *R. hominis*-specific primers 5'-CCCACTGACAGAGTATGTAATGTAC-3' and 5'-GCACCACCTGTCACCAC-3' were used for semi-quantitative and Real-time PCR analysis of faecal samples to validate gut colonization levels. Further bacterial PCR primers were designed using the on-line tool Primer3Plus (4) and purchased from Sigma-Aldrich. Real-time PCR analysis was performed using a 7500 Fast Real-Time PCR System (Applied Biosystems) with the Power SYBR Green PCR Master Mix (Applied Biosystems). PCR was performed as follows: one cycle at 95°C for 10 min, followed by 40 cycles at 95°C for 15 sec and 60°C for 1 min, ending with a dissociation step. All samples were run in triplicate. *GyrA* was used as a reference gene for normalization due to its low variation between samples.

For host gene expression, 2µg of total eukaryotic RNA isolated from the ileum and ascending colon was reverse-transcribed into cDNA using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) with random primers. Real-time PCR analysis was performed using a 7500 Fast Real-Time PCR System (Applied Biosystems) with the QuantiFast SYBR Green PCR Kit (Qiagen) and QuantiTect Primer Assays (Qiagen). PCR cycling conditions were as follows: one cycle at 95°C for 5 min, followed by 40 cycles at 95°C for 10 seconds and at 60°C for 30 seconds, ending with a dissociation step. All samples were run in triplicate. *Hprt* was selected as a reference gene for normalization because of its low variation between samples.

All RT-PCR data were analyzed on a logarithmic scale with base 2 by one-way ANOVA with a significance cut-off of $P < 0.05$. Differences were back-transformed to calculate fold changes.

Western blot

Immuno-purified rabbit polyclonal antibodies against *Roseburia hominis* Fla1 and Fla2 was produced as described in Duck *et al.* (5). In brief, New Zealand white female rabbits were immunized with synthetic peptide in complete Freund's adjuvant and boosted several times. For *R. hominis* Fla1, peptide NH₂-CRSQVRGLNKASDNA-CONH₂ and peptide NH₂-IDGNFTSKKLQVGSLC-COOH were used, while for *R. hominis* Fla2, peptide C-

AQYNDDAKSVLEILK-COOH and peptide C-GLNKASRNSQDGIS-CONH₂ were used. Following immunization the antibodies were purified on an immunoaffinity column prepared by coupling the peptides to 1 mL of activated sepharose beads.

For the western blot, *R. hominis* was grown in the presence of varying amounts (0.01g to 1g diet/10 mL of culture) of UV irradiated standard mouse chow for 3hrs, filtered to remove dietary components and diluted in laemmli buffer containing 8M urea. Thirty µL of each sample was loaded into wells of a NuPAGE® Novex® 4-12% Bis-Tris gel (Invitrogen) and electrophoresed, followed by further processing using the WesternBreeze Chromogenic Immunodetection System (Invitrogen). Fla1 and Fla2 antibodies were diluted 1:1000 and loading control anti-DNA gyrase A (Abcam) diluted 1:300 in antibody diluent and incubated overnight at 4°C, followed by 1h at room temperature with alkaline phosphatase conjugated anti-rabbit. Detection was by substrate colour development relative to loading control colour development.

FISH analysis

Tissues fixed in neutral buffered formalin were embedded in Technovit 8100 (Heraeus Kulzer). Two-micron sections were cut using a rotary microtome (Leica/Reichert Autocut). Three sections were taken per slide at 100 µm, 200 µm and 300 µm into the tissue, resulting in nine sections per animal.

Slides were dehydrated by consecutive incubations in 50% (v/v), 80% and 96% ethanol and dried at room temperature (RT). The 16S rRNA FISH probes used were a general bacterial probe Eub338 (GCTGCCTCCCGTAGGAGT; Cy3) and a newly designed *R. hominis* A2-183-specific probe (GTACATTACATACTCTGTCAAGTG; FITC), which was extensively tested for specificity against a panel of intestinal bacterial isolates. Ten microliter probe (30 ng/µL) in 100 µL hybridization buffer was applied to the dehydrated sample and incubated at probe-specific temperature. The slides were washed in washing buffer at 50°C for 30 min, dipped in ice-cold water to remove residual washing buffer and dried under compressed air flow. Counterstaining was performed with 4',6-diamidino-2-phenylindole (DAPI; Vector Laboratories Inc) and slides were mounted with Vectashield Mounting Medium for fluorescence (Vector Laboratories Inc) to prevent fading. Bacteria were visualized using a Leica DM RBE fluorescence microscope (Leitz GMBH) and photographed with a Penguin 600CL camera (Pixera) and Viewfinder 3.0 software (Studio Lite). High-magnification images (x630) were retrieved using the Apochromatics system (Leica).

Immunofluorescence

Immuno-localization of *R. hominis* flagellin was examined in colon contents of mice colonized with *R. hominis* using specific antisera raised against defined peptide sequences from both Fla1 and Fla2 flagellin proteins. Gut contents were diluted in PBS, smeared on glass slides and air dried. Smears were fixed in pre-cooled methanol for 5 min at -20 °C , incubated with anti-Fla1 or anti-Fla2 rabbit antisera (1:125, CovaLabs) overnight at 4 °C and visualized using Alexa donkey anti rabbit 488 (1:1000, Molecular Probes).

Sections were fixed in pre-cooled methanol for 30 min at -20 °C. Immuno-localization of T cell markers was examined on sequential cryosections (8µm). Sections were fixed either in pre-cooled methanol for 30 min at -20 °C (Ly6G FITC, CD3 FITC, CD11b FITC, all at 1:50 (BD Biosciences)), or, for the double-labelled FoxP3 (1:500, Abcam) with CD3 FITC (1:100, BD Biosciences) fixed in 1% paraformaldehyde (PFA) for 2 min at RT followed by 3 min in 0.01% Triton X in PBS. All sections were blocked with 10% BSA (Sigma) containing 10% relevant pre-immune sera in PBS (pH 7.4). Methanol-fixed tissues were incubated with primary antibodies for 1h at RT. PFA-fixed sections were incubated with antibodies overnight at 4 °C. FoxP3 was visualized using Alexa goat anti rabbit 594 (1:1000, Molecular Probes). Sections were counter labelled with DAPI and mounted with Vectashield (Vector Laboratories). For quantification of positive cells, a minimum of five fields of view from each mouse section was examined, using imaging software and microscope settings described above.

Histology

Ascending colon tissue samples were fixed for three hours in neutral buffered formalin (Sigma) at room temperature with constant agitation. The samples rinsed in PBS and then transferred to 70% ethanol and stored at room temperature until orientated for transverse sectioning and embedded in cold-curing resin using Technovit 8100 (Heraeus Kulzer) according to the manufacturer's instructions. The embedded tissue was mounted onto Histoblocs using Technovit 3040 (Heraeus Kulzer). Four micron sections were cut using a rotary microtome (Leica Autocut) fitted with a glass knife (TAAB Laboratories Equipment Ltd.). Tissue sections were stained using standard haemotoxylin/eosin methods. A complete transverse cross sectional area of an ascending colon from each animal was imaged at x200 magnification on a Zeiss Axioskop microscope using a QImaging camera controlled by Image Pro Plus software. Each field of view was then scored from 0-4 according to a method based on Berg *et al.* (6). Histopathology scores were 0 = Shallow crypts, no or few infiltrating inflammatory cells, intact epithelium, goblet cells appear full of mucin (no pathology); 1 = Crypts may exhibit slight epithelial cell hyperplasia, some diffuse infiltrating inflammatory cells may be seen between crypts, luminal epithelium appears intact, goblet cells may appear

slightly depleted of mucin; 2 = Crypts appear deeper with distinct evidence of epithelial hyperplasia, depletion of mucin from goblet cells, infiltrating inflammatory cells evident and may be multifocal in nature, although infiltrates are not seen in the submucosa; 3 = Lesions involve a larger area of the mucosa and /or are more frequent than seen in grade 2. Lesions do not involve the submucosa. Luminal epithelial cells exhibit small erosions. The lesions are not transmural; 4 = Crypt epithelium appears eroded. Abscesses may be present. Luminal epithelial cells appear irregular, sometimes with complete loss. Transmural infiltrate is observed - often associated with complete loss of epithelial cells into the lumen.

The mean percentage of fields of view at a given grade was calculated and treatment groups were compared using student t-test analysis.

Cloning and purification of recombinant flagellins

Flagellin genes were isolated from liquid bacterial cultures of *R. hominis*, *R. intestinalis*, *S. typhimurium*, *S. enteritidis*, *Eubacterium rectale* 33656 and *E. coli* K12 by PCR amplification and purification. The purified flagellin fragment from *R. hominis* was inserted into expression vector pT7-MAT-Tag-FLAG (Sigma) and flagellin fragments from *S. enteritidis* and *E. coli* K12 were inserted into expression vector pGEX-6P-1 (GE Healthcare). Recombinant flagellin was expressed by transformation of plasmid DNA into *E. coli* BL21 Rosetta and *E. coli* BL21 (DE3) cells, respectively and induction with 1mM IPTG (isopropyl b-D-galactosidase). Flagellins were recovered from cell lysates with FLAG beads (Sigma) and Ni-NTA (nickel-nitriloacetic) beads (Clontech, Takara) as per the manufacturer's instructions. The purity of the preparation was assessed by SDS-PAGE stained with coomassie blue solution. Activity of the flagellins fragments was determined by Luciferase Assay (Promega) according to the manufacturer's instructions using the NF- κ B transformed Caco-2 cell line.

Caco-2 cells were incubated with recombinant flagellins at a final concentration of 100 ng/ μ L for 2h at 37°C in a 75% humidified atmosphere of 5% CO₂. After the treatment the cells were washed with PBS solution twice and harvested for total RNA isolation.

Isolation of intestinal and MLN cells

Cells were isolated from the small intestine and mesenteric lymph node as previously described with minor modifications (7). Briefly, cellular suspensions were incubated with 100U/ml collagenase VIII (Sigma-Aldrich) in RPMI supplemented with 20% FBS at 37°C for 20min (mesenteric lymph nodes) or 1 hour (intestinal pieces). Single cell suspensions were then analyzed by flow cytometry.

Generation of bone marrow-derived dendritic cells and cultures

Bone marrow was harvested from femur and tibia of C3H/HeN and C57Bl6 mice as previously described (8-11). For GM-CSF-derived dendritic cells, bone marrow cells were resuspended at $1 \times 10^6/\text{mL}$ in RPMI supplemented with 10% FCS and 20ng/mL rmGM-CSF and seeded at 10mL/plate in 100mm² tissue culture plates. After three days culture, loosely adherent cells were collected and replated with GM-CSF supplemented media at $1 \times 10^6/\text{mL}$ in 12 well tissue culture plates. At day 5, cells were stimulated with 100ng/mL flagellins before being harvested on day 6. For Flt3L-derived dendritic cells, bone marrow cells were resuspended at $2 \times 10^6/\text{mL}$ in RPMI supplemented with 10% FCS and 200ng/mL rmFlt3 and seeded at 2mL/well in 12-well tissue culture plates. Cells were cultured for 10 days with an additional 2mL of Flt3 supplemented media added to each well on day 4. At day 9, cells were stimulated with 100ng/mL flagellins before being harvested on day 10. Cells were harvested from plates by gentle pipetting and analyzed by flow cytometry.

Flow Cytometry

Single cell suspensions of lamina propria cells, mesenteric lymph node cells and dendritic cells were incubated in blocking buffer (containing serum and CD16/CD32 antibody) at 4°C for 15 min prior to staining with specific fluorochrome conjugated antibodies. Lamina propria cells were labelled with antibodies to mouse CD4-FITC and CD25-APC (eBioscience), CD8-APC-Cy7, CD3-PerCP Cy5.5 and B220-BV570 (Biolegend). Intracellular FoxP3-PE (eBioscience) labelling was performed after extracellular staining and cell fixation/permeabilisation according to the manufacturer's instructions (eBioscience). GM-CSF derived dendritic cells were labelled with antibodies CD11b-PerCP Cy5.5 (BD Biosciences), CD11c-PE-Cy7, I-A/I-E -APC-Cy7, CD80-PE, CD86-APC, CD8-FITC, B220-BV570 (Biolegend). Flt3L derived dendritic cells were labelled with CD11c-PE-Cy7-, CD11b- or Siglec-H-PerCP Cy5.5 (Biolegend), I-A/I-E -APC-Cy7, CD317-PE, CD40-Alexa Fluor 647, CD103-FITC, B220-BV570. Cells were analyzed using a FACSAriaII (BD Biosciences) and FlowJo software version 7.2.5.

Cytometric Bead Array (CBA)

Bone marrow cells were isolated from femur and tibia with RPMI media of C3H/HeN and C57Bl/6 mice and Flt3L-expanded as described previously. Cells were stimulated with 100 ng/mL flagellin (e.g. *Roseburia* flagellin) after 9 days of culture, and supernatant was collected on day 10. The experiment was performed on three separate occasions to create N=3.

CBA analysis was performed on cell supernatants using the Cytometric Bead Array Mouse

Enhanced Sensitivity Master Buffer Kit (BD Biosciences) according to the manufacturer's instructions. Standards and samples were loaded onto a 96-well plate for measurement in a FACSArray (BD Biosciences). Results were analyzed using BD FCAP software (BD Biosciences).

Dry body weight and lipid carcass analysis

Eviscerated mouse carcass was weighed, lyophilized to constant weight and then milled for analysis. Lipid content was determined by extraction (1:100 w/v) with chloroform/methanol (2:1 v/v) as described previously (12).

Cloning of Roseburia flagellins.

Amplification by PCR of the flagellin gene from liquid bacteria culture.

Roseburia hominis A2-183 flagellin sequences were retrieved from the National Center for biotechnology Information (NCBI) website (<http://www.ncbi.nlm.nih.gov/>).

Roseburia intestinalis L1-82^T (NCIMB = 13810; DSM = 14610^T) flagellin sequences were retrieved from the National Center for biotechnology Information (NCBI) website (<http://www.ncbi.nlm.nih.gov/>). *Accession numbers for the flagellin genes and proteins are provided in Table B* (Table B – summary of the accession numbers for *Roseburia* flagellins described herein) ***

Total RNA isolation from bacteria

Total RNA was isolated from 1 mL of log phase bacteria liquid culture using the RNeasy minikit (Qiagen, Sussex, UK), coupled to RNase-free DNase I (Qiagen) digestion and based on manufacturer's protocol. In brief, one milliliter of liquid culture was placed in a 1.5 mL tube, centrifuged at 5.000 x g for 5 min at 4°C and the supernatant was discarded. The bacterial pellet was resuspended in 100 µL of lysozyme-containing TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) and incubated at RT (room temperature). Gram-positive bacteria RH (*Roseburia hominis*) was incubated for 10 min with a lysozyme concentration of 3 mg/mL. Three hundred and fifty microliters of lysis buffer containing 1 % β-mercapto-ethanol were then added to the tube to disrupt the cells by vortexing and 250 µL of 100 % ethanol (Molecular grade, Merck) were added. Seven hundreds microliters of the resulting suspension was applied to the RNeasy mini column fitted with a catching tube, centrifuged for 15 seconds at 8.000 x g and the catch emptied. Again, 350 µL of wash buffer was applied to the column, which was centrifuged for 15 s at 8.000 x g and the flow-through discarded. DNA digestion was carried out by adding to each column, 30 units of DNase I mixed with 70 µL of RDD buffer, and incubating at RT for 15 min. The column was washed again as

Roseburia intestinalis (Fla3)

Forward primer (5'-3'): CTCGAGATATGGTAGTACAGCACAA

Reverse primer (5'-3'): CTTAGATCTCTGTAACAGAGAAAAGTA

Roseburia intestinalis (Fla4)

Forward primer (5'-3'): CCGGGATCCATGGTAGTACAGCACAAT

Reverse primer (5'-3'): TTAGTGGTGGTGATGATGATGCTGTAACAGAGAAAAG

Using 1 µL of bacterial culture as a template, the flagellin genes were amplified by PCR. The thermal program for amplification was adapted from Fermentas (Fermentas GMBH, Germany) for optimum KOD polymerase (Novagen, Madison, WI) functions and was 95°C for 2 min for polymerase activation followed by 34 cycles of amplification, 95°C for 20 s, annealing temperature for 10s and 70°C for 20 s ending with 2 min at 70°C using KOD polymerase. The details of the PCR mix are given in Table 1. The PCR products were run on a 1% agarose gel for 30 min at 120 V and the bands of interest were excised and purified.

The insert was first cloned into the cloning vector pCR™-Blunt II-TOPO using Zero Blunt TOPO PCR cloning kit (Invitrogen) according to the manufacturer's instructions and by mixing as described in Figure 1A and Table 1.

Table 1 Composition of the mix for the insertion of PCR product into pCR-Blunt II-TOPO cloning vector. Salt solution and vector-enzyme mix were ready-to-use solutions.

Components	Volume (µL)
PCR product	2
Salt solution	1
H ₂ O	2
Vector-enzyme mix	1
Total	6

Electrophoretic separation and purification of PCR products.

In order to view the isolated flagellin product amplified by PCR, the PCR reaction was mixed with 1 X Blue-orange loading dye (Promega) and loaded onto a 1% agarose gel (made of agarose 1% diluted in TAE buffer (40 mM Tris base, 0.1% (v/v) acetic acid, 1 mM EDTA) boiled and poured in a gel tray) containing 0.5 µg/mL of Ethidium bromide to stain nucleic acids. Samples migrated for 30 min under 120 V using TAE as running buffer. When migration was finished, the gel was visualized under ultra-violet (UV) and the product was identified using a reference standard and excised with a clean scalpel. The gel slice was then transferred into a pre-weighed eppendorf tube and its weight recorded by subtraction from the empty tube weight. The PCR product was purified. Membrane binding solution was added to the gel slice in a ratio of 10 µL of solution per 10 mg of agarose gel slice. The

mixture was then incubated at 65°C for 10 min with frequent vortexing until the gel slice was completely dissolved. The tube was centrifuged briefly at RT to ensure all contents were at the bottom of the tube before progressing to DNA purification. From this point onward the purification PCR involved using Wizard® SV Gel and PCR Clean-up system (Promega, Southampton, UK) according to the manufacturer's instructions. One SV minicolumn per PCR product was placed into a collection tube. The PCR product were added to the SV minicolumn assembly and incubated for 1 min at room temperature (RT). The column was centrifuged at 16.000 x g for 1 min, and the flow-through discarded. The column was then washed by adding 700 µL of membrane wash solution and centrifuged for 5 min at RT. The flow-through was discarded and the assembly spun again for 1 min and dried at RT for 2 min to evaporate any residual ethanol. The column was then transferred into a fresh 1.5 mL nuclease-free tube. To recover the PCR product, 25 µL of nuclease-free water was applied to the centre of the column and DNA was eluted by 1 min incubation at RT and 1 min spin at 16.000 x g. The concentration of the eluate was measured spectrophotometrically using Nanodrop (Thermo Fisher Scientific) and stored at -20°C.

Digestion of the flagellin and expression vector.

Two hundred microliters of freshly thawed DH5α competent cells were kept on ice and 4 µL of the PCR product mix, and left on ice for 5 min. Heat shock was performed for 1 min at 42°C and the cells returned on ice or 5 min, before addition of 400 µL of SOC medium. SOC Medium (Sigma) is a rich media used primarily in the recovery step of *Escherichia coli* competent cell transformations. Use of SOC maximizes the transformation efficiency of competent cells.

The cells grew at 37°C for 1 hour, and 200 µL of the mix was plated on LB agar plate supplemented with Kanamycin 50 µg/mL. The positive clones were prepared as minipreps, nanodropped and digested with BamH1 and Xho1 for 2 hours to excise the flagellin. The expression vector pT7-MAT-Tag-FLAG-2 (Sigma), depicted in Figure 2A, was also digested with BamH1 and Xho1. After separation of the product on a 1% agarose gel and purification of the digested vector and fragment i.e. the vector and the insert were ligated at 4°C overnight and transformed into *E.coli* DH5a. Ligation products were plated on LB agar supplemented with ampicillin 100 µg/mL. Positives clones were prepared as minipreps and transformed into *E.coli* BL21 (DE3) competent cells. The positives colonies were sequenced with appropriate primers to check the presence of the insert. Glycerol aliquots were made out the positive clones. The expression vectors encoding *Roseburia* flagellins were further prepared as minipreps and transformed into *E.coli* competent cells to improve the yield of the purification; for Rh1, Rh2 and Ri2 the *E. coli* competent cells were *E.coli* BL21 Rosetta and

for RI3, Ri1 Ri4, Se, St, K12, and Er the *E. coli* competent cells were *E. coli* BL21 (DE3).

Expression and purification of Roseburia flagellins

The transformed *E. coli* (e.g. BL21 Rosetta) encoding the flagellin gene (RH1 or FLaA1) were grown in 1 liter LB medium with ampicillin (100 µg/mL), and supplemented with chloramphenicol (50 µg/mL) for Rosetta cells, at 37°C under shaking 180 rpm until log phase. The culture was induced for 3 hours with 1 mM IPTG under the same conditions, to allow the specific expression of flagellin protein. Bacteria were pelleted by centrifugation 4.000 x g, 4°C, for 10 min and resuspended in 60 mL of buffer 1 (50 mM Tris (pH 7.4), 150 mM NaCl) and sonicated with a Soniprep 150 (Sanyo, Japan), for 1 min when using BL21 (DE3), or 3 min when using Rosetta with 1 min rest after each minute. Cell lysate was centrifuged at 9.000 x g for 10 min at 4°C and the supernatant discarded. The pellet, which contains insoluble proteins was resuspended in 60 mL of buffer 2 (50 mM Tris (pH 8.0), 100 mM NaCl, 5 mM EDTA, 0.5% Triton-X100, 1 mM DTT) sonicated again until complete resuspension, and centrifuged at 5.000 x g for 15 min at 4°C. The insoluble fraction was washed a second and a third time with buffer 2. A last wash was carried out using buffer 2 without Tx-100 and DTT, as described above. To solubilise the proteins, the insoluble fraction was resuspended using sonication in 4 mL of urea buffer (2 M urea, 50 mM Tris (pH 7.4), 150 mM NaCl) and centrifuged at 5.000 x g, 5 min at RT. The supernatant was discarded by decanting and the pellet was resuspended in 2 mL of 4 M urea, 50 mM Tris (pH7.4), 150 mM NaCl by sonication and centrifuged at 5.000 x g, 5 min at RT. The supernatant was collected by decanting and the pellet was resuspended in 2 mL of 8 M urea, 50 mM Tris (pH7.4), 150 mM NaCl solution by sonication and centrifuged at 5.000 x g, 5 min at RT. The supernatant was collected by decanting and mixed to the 4M urea fraction, to make a 6M urea solution containing solubilized flagellins. One milliliter of Talon resin (Clontech, Takara, UK) was washed twice with 2 mL of buffer 1 (50 mM Tris (pH7.4), 150 mM NaCl) and centrifugated at 1.000 x g, 2 min at RT. The supernatant was discarded and the 6 M urea fraction containing flagellin was added, and left for rotation at 4°C for 1 hour. After centrifugation at 1.000 x g for 2 min at RT, the resin was washed twice with buffer containing 2 M urea, 50 mM Tris (pH 7.4), 150 mM NaCl, 15 mM imidazole. After centrifugation, the resin was incubated for 10 min with 300 µL of elution buffer (360 mM Tris-HCl, 680 mM imidazole, 360 mM NaCl, 0,35 N HCl, 1.43 M urea), gently mixed, and the eluate was collected after centrifugation at 2.000 x g for 2 min at RT. The elution was repeated once and the two eluates mixed. A second purification step using anti-FLAG M2 magnetic beads (Sigma) was performed to improve the purity of the protein preparation. Four hundred microliters of ultra pure water were slowly added to the eluate of the Talon resin, to

obtain a final concentration of 0.9 M urea, suitable for FLAG beads. Two hundred fifty microliters of FLAG magnetic beads were washed twice with 750 μ L of buffer 1, by mixing and leaving on a magnetic stand (Ambion) to remove the supernatant. The Talon-eluate was added to the beads and rotated for 30 min at 4°C. The beads were washed five times with 1 mL of 0.9 M urea, 50 mM Tris (pH7.4), 150 mM NaCl. The elution consisted in adding to the beads 250 μ L of 200 μ g/mL FLAG peptide solution containing 0.9 M urea, 50 mM Tris (pH7.4), 150 mM NaCl, mixing well and incubating 30 min at RT. The tubes were placed on the magnetic stand and the final product was collected. This elution step was repeated once and 500 μ L final products were further dialyzed in PBS with 0.9 M urea for 1 hour at 4°C, prepared as aliquots and stored at -80°C. The purity of the preparations was assessed by SDS-PAGE stained with coomassie blue solution (Figure 3A). Just before use, the protein aliquot was dialyzed 1 hour against PBS, changing the buffer twice, in slide-a-lyser of 20 kDa molecular weight cut off (Pierce, UK) and the protein concentration was measured using Bradford's method. Endotoxin measurement confirmed values were <0.25 EU/mL.

The nucleotide sequence encoding Fla1 is shown in SEQ ID NO 1 and the amino acid sequence of Fla1 is shown in SEQ ID NO 2.

ATGGTAGTACAGCACAATCTTACAGCAATGAACGCTAACAGACAGTTAGGTATCACAAC
 AGGCGCACAGGCTAAGTCTTCTGAGAAGTTATCTTCTGGTTACAAGATCAACCGCGCAG
 CAGATGACGCAGCAGGTCTTACGATTTCCGAGAAGATGAGAAGCCAGGTTAGAGGCTT
 AAATAAAGCTTCTGACAACGCACAGGATGGTGTATCCCTTATTCAGGTAGCTGAGGGTG
 CATTAAGTGAGACACACTCCATCTTACAGCGTATGAATGAGTTAGCAACTCAGGCAGCA
 AACGATACCAATACAACCTCAGACAGAACTGCAGTTCAGCAGGAGATCAACCAGTTAGC
 ATCTGAGATCACCAGAATCGCTTCTACAACTCAGTTCAACACAATGAACCTGATCGATGG
 TAACTTCACAAGTAAGAAGCTTCAGGTAGGTTCCCTTTGCGGACAGGCTATCACAATCG
 ATATCTCTGATATGTCTGCTACAGGTCTTGGCGTTAGCGGATTAGTAGTATCTTCCTTCT
 CTGCAGCTGGTAAGGCAATGTCTGCAGCTCAGGATGCTATCAGCTACGTATCTTCTATG
 CGTTCTAAGCTGGGTGCATTACAGAACAGACTTGAGCACACAATCTCCACCTGGACAA
 CATTTCTGAGCACACATCTTCTGCAGAGTCTCGTATCCGTGATACAGATATGGCTGAAG
 AGATGGTTGAGTACAGCAAGAACAACATCCTTGCTCAGGCAGGACAGTCTATGCTTGCT
 CAGGCTAACCAGTCTACTCAGGGTGTATTATCCTTATTACAGTAA (SEQ ID NO 1).

MVVQHNLTAMNANRQLGITTGAQAKSSEKLSSGYKINRAADDAAGLTISEKMRSQVRGLNK
 ASDNAQDGVSLIQVAEGALSETHSILQRMNELATQAANDTNTTSDRTAVQQEINQLASEITRI
 ASTTQFNTMNLIDGNFTSKKLQVGS LCGQAITIDISDMSATGLGVSGLVVSSFSAAGKAMSA
 AQDAISYVSSMRSLGALQNRLEHTISHLDNISEHTSSAESRIRDTDMAEEMVEYSKNNILAQ

AGQSMLAQANQSTQGVLSLLQ (SEQ ID NO 2).

The flagellin structure consists of four domain D0, D1, D2 and D3. See Figure 4A.

D0: N-terminal α -helix starts from Gln 2 and extends up to Ser 32 (ND0). C-terminal α -helix starts from Ala 459 and extends down to Ser 491 (CD0).

The spoke region, which connects D0 and D1 domains, consists of two chains (Ns and Cs), one from Ser 32 to Ala 44 and the other from Glu 454 to Ala 459.

D1: N-terminal segment extends from Ala 44 to Gln 176 and the C-terminal segment from Asn 406 to Glu 454. N-terminal segment is made of an α -helix (from Ala 44 to Ala 99) (ND1a), followed up by a loop connecting to the second, shorter α -helix (ND1b) which goes down and the chain continues to two β -turns, a β -hairpin pointing down and an extended chain going up, and the rest of the chain finally goes into domain D2. The C-terminal α -helix in domain 1 (CD1) starts from Asn 406 and extends to Glu 454.

Domain D0 and D1 are packed into the protofilament structure so that N-D0 is oriented to the outside and C-D0 exposed to the central channel.

D2 domain comprises two segments: N-terminal segment from Lys 177 to Gly 189 and C-terminal segment from Ala 284 to Glu 405. It is made mostly of β strands excepting two helix 285-289 and 288-298.

D3 domain comprises a central segment from Tyr 190 to Val 283. Mostly made of β strands with one short stretch of helical fold (199-209).

Without wishing to be bound by theory, two essential regions of flagellin protein involved in the recognition and activation of TLR5 are thought to be amino acids at positions 79-117 of SEQ ID NO 2 (N-D1 domain) and amino acids at positions 408-439 of SEQ ID NO 2 (C- D1 domain), as well as the amino acid alanine (A) at position 411 of SEQ ID NO 2, the amino acid glutamine (Q) at position 412 of SEQ ID NO 2 and the amino acid serine (S) at position 420 of SEQ ID NO 2.

EXAMPLES

***R. hominis* preferentially colonizes the colon**

C3H/HeN germfree (GF) mice were inoculated with three gavages of *R. hominis* on consecutive days. The inventors report the first successful mono-colonization of germfree mice with a single bacterial species from the Firmicutes phylum. Successful colonization was achieved using an inoculation medium containing 3% ascorbic acid and 2% cysteine to protect this oxygen sensitive bacterium. Analysis of gut tissue by fluorescent *in situ*

hybridization (FISH) revealed that *R. hominis* colonized both the ileum and colon, but was found in much higher numbers in the colon. Bacteria were also found closely associated with the colonic mucosa (Fig. S1A). Colonization was further validated and quantified by PCR using *R. hominis*-specific primers with numbers approximating 1×10^{10} bacteria/g faeces (Fig. S1B-C). Faeces of GF animals tested negative for the presence of any bacteria.

The *R. hominis* genome reveals unique genes promoting host interactions

The complete genome sequence of *R. hominis* A2-183 was generated and is represented by a single 3,592,125-bp chromosome (Fig. 1A). Automated and manual annotation of the genome using the RAST platform revealed the presence of four ribosomal operons, 66 RNAs and 3,273 predicted proteins. The largest group of genes belonged to the Subsystem Category Carbohydrates (271 genes), encoding proteins involved in carbohydrate metabolism, followed by Protein Metabolism (197) and Amino acids and Derivatives (175) (Fig. 1B). Other important functional categories included Motility and Chemotaxis (49) and Dormancy and Sporulation (12). Comparative genomic analysis established that the closest relative in terms of genomic structure and function among the complete bacterial genomes is *Eubacterium rectale* (Mahowald *et al.* 2009), which is not surprising given the close taxonomical relatedness of these organisms (Duncan *et al.* 2006, Aminov *et al.* 2006). Comparative reconstruction of the two genomes with 1,095 genes revealed that they differed by approximately 25% of the genes. In particular, these differences encompassed genes encoding important functions for interaction with the host. For example, the Motility and Chemotaxis genes encoding type IV fimbrial assembly proteins PilB and PilC were present in *E. rectale* but absent in *R. hominis* whereas flagellar basal-body rod protein FlgC, flagellar hook-basal body complex protein FliE, flagellin protein FlaB and flagellar motor switch protein FliG were unique to *R. hominis* (Table S1). The two bacterial genomes also differed by 42 carbohydrate genes, reflecting their divergent nutritional requirements. Interestingly, *R. hominis* is unique in that it expresses two flagellin genes that are both FLA type as opposed to flagellins homologous to FliC expressed by Salmonella and *E. coli* bacterial species.

Table S1. Comparative analysis of the *R. hominis* and *E. rectale* genomes.

Comparative genomic analysis showed that *Eubacterium rectale* is the closest known relative of *R. hominis* in terms of genomic structure and function, and the two genomes differed by approximately 25% of the genes. (A) Genes present in *R. hominis* but absent in *E. rectale* and (B) genes present in *E. rectale* but absent in *R. hominis*.

Table S1A. Genes present in *R. hominis* but absent in *E. rectale*.

Subcategory	Role
Amino Acids and Derivatives	
Alanine, serine, and glycine	L-serine dehydratase, alpha subunit
Alanine, serine, and glycine	L-serine dehydratase, beta subunit
Arginine; urea cycle, polyamines	Arginine pathway regulatory protein ArgR, repressor of arg
Arginine; urea cycle, polyamines	NADP-specific glutamate dehydrogenase
Arginine; urea cycle, polyamines	Ornithine decarboxylase
Arginine; urea cycle, polyamines	Spermidine Putrescine ABC transporter permease
Aromatic amino acids and derivatives	Biosynthetic Aromatic amino acid aminotransferase alpha
Branched-chain amino acids	Threonine dehydratase
Glutamine, glutamate, aspartate, asparagines	Aspartate aminotransferase
Glutamine, glutamate, aspartate, asparagine	L-asparaginase
Lysine, threonine, methionine, and cysteine	Sulfate permease
Lysine, threonine, methionine, and cysteine	Lysine decarboxylase
Lysine, threonine, methionine, and cysteine	Adenosylhomocysteinase
Lysine, threonine, methionine, and cysteine	Methionine ABC transporter ATP-binding protein
Lysine, threonine, methionine, and cysteine	Methionine ABC transporter permease protein
Lysine, threonine, methionine, and cysteine	Methionine ABC transporter substrate-binding protein
Lysine, threonine, methionine, and cysteine	S-ribosylhomocysteine lyase
Lysine, threonine, methionine, and cysteine	Predicted functional analog of homoserine kinase
Carbohydrates	
Aminosugars	Beta-hexosaminidase
Central carbohydrate metabolism	6-phosphogluconolactonase
Central carbohydrate metabolism	Aldose 1-epimerase
Central carbohydrate metabolism	4-Hydroxy-2-oxoglutarate aldolase
Central carbohydrate metabolism	Pyruvate:ferredoxin oxidoreductase, alpha subunit
Central carbohydrate metabolism	Pyruvate:ferredoxin oxidoreductase, beta subunit
Central carbohydrate metabolism	Pyruvate:ferredoxin oxidoreductase, delta subunit
Central carbohydrate metabolism	Pyruvate:ferredoxin oxidoreductase, gamma subunit
Central carbohydrate metabolism	Malate permease
Di- and oligosaccharides	Beta-glucoside bgl operon antiterminator, BglG family
Di- and oligosaccharides	Galactose operon repressor, GalR-LacI family of
Di- and oligosaccharides	Galactose/methyl galactoside ABC transport system, permease protein MglC

Di- and oligosaccharides	Multiple sugar ABC transporter, ATP-binding protein
Fermentation	Enoyl-CoA hydratase
Fermentation	D-lactate dehydrogenase
Monosaccharides	2-dehydro-3-deoxygluconate kinase
Monosaccharides	2-dehydro-3-deoxyphosphogluconate aldolase
Monosaccharides	2-deoxy-D-gluconate 3-dehydrogenase
Monosaccharides	Alpha-glucosidase
Monosaccharides	Altronate hydrolases
Monosaccharides	Altronate oxidoreductase
Monosaccharides	Beta-glucuronidase
Monosaccharides	D-mannonate oxidoreductase
Monosaccharides	Mannonate dehydratase
Monosaccharides	Rhamnogalacturonides degradation protein RhiN
Monosaccharides	Uronate isomerase
Monosaccharides	Sugar diacid utilization regulator SdaR
Monosaccharides	Ribose ABC transport system, ATP-binding protein RbsA
Monosaccharides	Ribose ABC transport system, permease protein RbsC
Monosaccharides	Pyrimidine-nucleoside phosphorylase
Monosaccharides	Fructokinase
Monosaccharides	Transcriptional repressor of the fructose operon, DeoR
Monosaccharides	Mannose-1-phosphate guanylyltransferase
Monosaccharides	Mannose-6-phosphate isomerase
Monosaccharides	Possible alpha-xyloside ABC transporter, substrate-binding
One-carbon Metabolism	Serine-pyruvate aminotransferase
Sugar alcohols	Glycerol uptake facilitator protein
Sugar alcohols	Glycerol-3-phosphate ABC transporter, periplasmic glycerol-3-phosphate-binding protein
No subcategory	Carbon storage regulator
Cell Wall and Capsule	
Capsular and extracellular polysaccharides	D,D-heptose 7-phosphate kinase
Capsular and extracellular polysaccharides	GDP-L-fucose synthetase
Capsular and extracellular polysaccharides	dTDP-4-dehydrorhamnose 3,5-epimerase
Capsular and extracellular polysaccharides	Peptidoglycan N-acetylglucosamine deacetylase
Capsular and extracellular polysaccharides	Polysaccharide deacetylase

Capsular and extracellular polysaccharides	Glucose-1-phosphate thymidyltransferase
Capsular and extracellular polysaccharides	Teichoic acid translocation permease protein TagG
Capsular and extracellular polysaccharides	dTDP-4-dehydrorhamnose reductase
Capsular and extracellular polysaccharides	dTDP-glucose 4,6-dehydratase
Capsular and extracellular polysaccharides	TRAP-type transport system, small permease component, predicted N-acetylneuraminase transporter
Capsular and extracellular polysaccharides	Glucose-1-phosphate cytidyltransferase
No subcategory	N-acetylmuramoyl-L-alanine amidase
Clustering-based subsystems	
Chemotaxis, response regulators	diguanylate cyclase (GGDEF domain) with PAS/PAC
Clustering-based subsystems	Stage II sporulation protein related to metalloproteases
Cytochrome biogenesis	Glutamate-1-semialdehyde aminotransferase
Protein export?	Membrane proteins related to metalloendopeptidases
Protein export?	NLP/P60 family protein
Protein export?	Peptide chain release factor 2
Protein export?	hypothetical protein BH3604
No subcategory	Cell division protein FtsW
No subcategory	FIG003307: hypothetical protein
No subcategory	SSU ribosomal protein S1p
Cofactors, Vitamins, Prosthetic Groups, Pigments	
Coenzyme A	2-dehydropanoate 2-reductase
Coenzyme A	Pantothenate:Na ⁺ symporter
Fe-S clusters	Iron-sulfur cluster assembly protein SufD
NAD and NADP	L-Aspartate dehydrogenase
Pyridoxine	Predicted transcriptional regulator of pyridoxine metabolism
Tetrapyrroles	Alpha-ribazole-5'-phosphate phosphatase
No subcategory	Predicted hydroxymethylpyrimidine transporter CytX
DNA Metabolism	
CRISPs	CRISPR-associated protein Cas2
DNA repair	Uracil-DNA glycosylase, family 1
DNA repair	Excinuclease ABC subunit A paralog of unknown function
DNA repair	A/G-specific adenine glycosylase
DNA repair	Exodeoxyribonuclease VII small subunit
DNA repair	Exonuclease SbcC

DNA repair	Exonuclease SbcD
DNA repair	putative ATP-dependent DNA helicase YjcD
DNA replication	DNA polymerase III subunits gamma and tau
DNA replication	DNA replication protein DnaC
No subcategory	DNA-binding protein HBsu
Dormancy and Sporulation	
Spore DNA protection	Small acid-soluble spore protein, beta-type SASP
No subcategory	Spore maturation protein A
No subcategory	Spore maturation protein B
Fatty Acids, Lipids, and Isoprenoids	
Fatty acids	4'-phosphopantetheinyl transferase
Triacylglycerols	Lysophospholipase
Triacylglycerols	Lysophospholipase L2
Triacylglycerols	Monoglyceride lipase
Membrane Transport	
ABC transporters	Phosphonate ABC transporter ATP-binding protein
ABC transporters	Phosphonate ABC transporter permease protein phnE1
ABC transporters	Phosphonate ABC transporter permease protein phnE2
ABC transporters	Dipeptide transport ATP-binding protein DppD
ABC transporters	Dipeptide transport ATP-binding protein DppF
ABC transporters	Dipeptide transport system permease protein DppB
ABC transporters	Dipeptide transport system permease protein DppC
ABC transporters	Dipeptide-binding ABC transporter, periplasmic substrate-
No subcategory	Duplicated ATPase component BL0693 of energizing
No subcategory	Duplicated ATPase component MtsB of energizing module of methionine-regulated ECF transporter
No subcategory	Substrate-specific component BL0695 of predicted ECF
No subcategory	Substrate-specific component MtsA of methionine-regulated
No subcategory	Substrate-specific component PdxU2 of predicted pyridoxin-
No subcategory	Substrate-specific component ThiT of thiamin ECF
No subcategory	Transmembrane component BL0694 of energizing module
No subcategory	Transmembrane component MtsC of energizing module of
Miscellaneous	
Plant-Prokaryote DOE project	Aspartyl-tRNA(Asn) amidotransferase subunit A

Plant-Prokaryote DOE project	Aspartyl-tRNA(Asn) amidotransferase subunit B
Plant-Prokaryote DOE project	Glutamyl-tRNA(Gln) amidotransferase subunit A
Plant-Prokaryote DOE project	Glutamyl-tRNA(Gln) amidotransferase subunit B
Plant-Prokaryote DOE project	Phosphoglucosamine mutase
Plant-Prokaryote DOE project	Phosphomannomutase
Plant-Prokaryote DOE project	ABC transport system, sugar-binding protein
Plant-Prokaryote DOE project	Alpha-L-arabinofuranosidase II precursor
Plant-Prokaryote DOE project	Alpha-N-arabinofuranosidase
Plant-Prokaryote DOE project	Alpha-N-arabinofuranosidase 2
Plant-Prokaryote DOE project	COG3533 secreted protein
Plant-Prokaryote DOE project	L-arabinose isomerase
Plant-Prokaryote DOE project	Transcriptional repressor of the arabinose operon
Plant-Prokaryote DOE project	rhamnogalacturonan acetyltransferase
No subcategory	Putative activity regulator of membrane protease YbbK
Motility and Chemotaxis	
Flagellar motility in Prokaryota	Flagellar basal-body rod protein FlgC
Flagellar motility in Prokaryota	Flagellar hook-basal body complex protein FlIE
Flagellar motility in Prokaryota	Flagellin protein FlaB
No subcategory	Flagellar motor switch protein FlgG
Nitrogen Metabolism	
No subcategory	Nitrogen regulatory protein P-II
No subcategory	Hcp transcriptional regulator HcpR (Crp/Fnr family)
Nucleosides and Nucleotides	
Purines	Adenine deaminase
Pyrimidines	Uridine kinase
No subcategory	Ribonucleotide reductase of class III (anaerobic), large
Phosphorus Metabolism	
No subcategory	Exopolyphosphatase
No subcategory	Phosphate transport system permease protein PstA
No subcategory	Phosphate transport system permease protein PstC
No subcategory	Sodium-dependent phosphate transporter
Potassium metabolism	
No subcategory	Potassium voltage-gated channel subfamily KQT
Protein Metabolism	

Protein biosynthesis		Probable GTPase related to EngC
Protein biosynthesis		Aspartyl-tRNA(Asn) synthetase
Protein biosynthesis		tRNA-Ala-CGC
Protein biosynthesis		tRNA-Gly-CCC
Protein biosynthesis		tRNA-Pro-GGG
Protein biosynthesis		tRNA-Ser-CGA
Protein biosynthesis		tRNA-Ser-GGA
Protein biosynthesis		tRNA-Val-CAC
Protein folding		Foldase protein PrsA precursor
Protein processing and modification		[NiFe] hydrogenase nickel incorporation protein HypA
Protein processing and modification		[NiFe] hydrogenase nickel incorporation-associated protein
RNA Metabolism		
RNA processing and modification		Peptidyl-prolyl cis-trans isomerase
RNA processing and modification		Ribonuclease P protein component
Transcription		Predicted transcriptional regulator of cysteine synthase,
Transcription		RNA polymerase sigma factor SigV
Regulation and Cell signaling		
Programmed Cell Death and Toxin-antitoxin Systems		YafQ toxin protein
No subcategory		Bifunctional autolysin Atl
No subcategory		Cell envelope-associated transcriptional attenuator LytR-
No subcategory		Aromatic hydrocarbon utilization transcriptional regulator
No subcategory		Catabolite control protein A
No subcategory		Catabolite repression HPr-like protein Crh
Respiration		
No subcategory		Ferredoxin
Stress Response		
Cold shock		Cold shock protein CspG
Virulence, Disease and Defense		
Resistance to antibiotics and toxic		Probable Co/Zn/Cd efflux system membrane fusion protein
Resistance to antibiotics and toxic		Transcriptional regulator, MerR family
Resistance to antibiotics and toxic		Heavy metal-(Cd/Co/Hg/Pb/Zn)-translocating P-type
Resistance to antibiotics and toxic		Vancomycin B-type resistance protein VanW
Resistance to antibiotics and toxic		Tetracycline resistance protein TetW

Table S1B. Genes present in *E. rectale* but absent in *R. hominis*.

Subcategory	Role
Amino Acids and Derivatives	
Arginine; urea cycle, polyamines	Transcriptional regulator, MerR family, near polyamine
Lysine, threonine, methionine, and cysteine	Sulfate adenylyltransferase subunit 2
Lysine, threonine, methionine, and cysteine	Sulfate and thiosulfate import ATP-binding protein CysA
Lysine, threonine, methionine, and cysteine	Sulfate transport system permease protein CysW
Lysine, threonine, methionine, and cysteine	Methionine transporter MetT
Carbohydrates	
Aminosugars	N-Acetyl-D-glucosamine ABC transport system,
Di- and oligosaccharides	Galactose/methyl galactoside ABC transport system,
Di- and oligosaccharides	Maltodextrin glucosidase
Di- and oligosaccharides	PTS system, maltose and glucose-specific IIB
Di- and oligosaccharides	PTS system, maltose and glucose-specific IIC
Fermentation	NADH-dependent butanol dehydrogenase A
Fermentation	Alcohol dehydrogenase
Monosaccharides	Predicted beta-xyloside ABC transporter, substrate-
One-carbon Metabolism	Fumarate hydratase class I, aerobic
Organic acids	Serine--glyoxylate aminotransferase
Polysaccharides	Glycogen biosynthesis protein GlgD, glucose-1-phosphate adenylyltransferase family
Polysaccharides	glycogen debranching enzyme-related protein
Sugar alcohols	Glycerol dehydrogenase
Cell Wall and Capsule	
Gram-Positive cell wall components	Cell wall surface anchor family protein
Clustering-based subsystems	
Cell Division	FIG001960: FtsZ-interacting protein related to cell division
Isoprenoid/cell wall biosynthesis: predicted undecaprenyl diphosphate phosphatase	penicillin-binding protein, putative
Probably organic hydroperoxide resistance related hypothetical protein	Homoserine kinase
No subcategory	Cell division topological specificity factor MinE
No subcategory	Low molecular weight protein tyrosine phosphatase
Cofactors, Vitamins, Prosthetic Groups, Pigments	
Biotin	3-ketoacyl-CoA thiolase

Biotin	Biotin synthase
Pyridoxine	4-hydroxythreonine-4-phosphate dehydrogenase
Pyridoxine	Pyridoxal kinase
Pyridoxine	Pyridoxine biosynthesis glutamine amidotransferase,
Pyridoxine	Pyridoxine biosynthesis glutamine amidotransferase,
Tetrapyrroles	Cob(I)alamin adenosyltransferase
Tetrapyrroles	Uroporphyrinogen-III methyltransferase
Tetrapyrroles	Vitamin B12 ABC transporter, B12-binding component
No subcategory	Sulfur carrier protein adenylyltransferase ThiF
No subcategory	Thiazole biosynthesis protein ThiG
No subcategory	Thiazole biosynthesis protein ThiH
DNA Metabolism	
CRISPs	CRISPR-associated protein Cas7
DNA repair	Recombinational DNA repair protein RecT (prophage
DNA replication	ATP-dependent DNA helicase RecQ
Fatty Acids, Lipids, and Isoprenoids	
Phospholipids	CDP-diacylglycerol--serine O-phosphatidyltransferase
Phospholipids	Diacylglycerol kinase
Iron acquisition and metabolism	
No subcategory	Sortase A, LPXTG specific
Membrane Transport	
No subcategory	Substrate-specific component FolT of folate ECF transporter
Motility and Chemotaxis	
Social motility and nonflagellar swimming in bacteria	Type IV fimbrial assembly protein PilC
Social motility and nonflagellar swimming in bacteria	Type IV fimbrial assembly, ATPase PilB
Nucleosides and Nucleotides	
Detoxification	Mutator mutT protein (7,8-dihydro-8-oxoguanine-
Purines	Uracil-xanthine permease
Purines	Adenine phosphoribosyltransferase
Purines	Adenosine deaminase
Pyrimidines	Uridine phosphorylase
Phages, Prophages, Transposable elements, Plasmids	
Phages, Prophages	Phage tail length tape-measure protein
Protein Metabolism	
Protein biosynthesis	Similar to ribosomal large subunit pseudouridine synthase D, Bacillus subtilis YjbO type

Protein biosynthesis	tRNA (Guanine37-N1) -methyltransferase
Protein biosynthesis	Alanyl-tRNA synthetase family protein
Protein degradation	Aminopeptidase C
Protein degradation	Deblocking aminopeptidase
RNA Metabolism	
RNA processing and modification	COG1720: Uncharacterized conserved protein
RNA processing and modification	tRNA(Ile)-lysidine synthetase
Respiration	
Electron accepting reactions	Adenylylsulfate reductase alpha-subunit
Electron accepting reactions	Adenylylsulfate reductase beta-subunit
Electron accepting reactions	Dissimilatory sulfite reductase (desulfoviridin), alpha and
Stress Response	
Osmotic stress	Choline binding protein A
Osmotic stress	Sarcosine oxidase alpha subunit
No subcategory	Carbon starvation protein A
No subcategory	transcriptional regulator, PemK family
Sulfur Metabolism	
No subcategory	Beta-galactosidase large subunit
No subcategory	Beta-galactosidase small subunit
Virulence, Disease and Defense	
Resistance to antibiotics and toxic	Aminoglycoside N6'-acetyltransferase
Resistance to antibiotics and toxic	Cytoplasmic copper homeostasis protein cutC

***R. hominis* responds to the gut environment by up-regulating motility, mobilization and chemotaxis genes**

To determine the genes differentially expressed by *R. hominis* in response to association with the host and diet, a microarray was constructed using 6,000 PCR fragments from the small-insert-size sequencing library. Subsequent Real-time PCR validation was performed on 42 differentially expressed genes (Tables S2 and S3), which cluster at specific regions of the *R. hominis* genome as illustrated in Fig. 2A. To distinguish between the effects of gut environment and dietary components, bacterial RNA was isolated from four different experimental conditions: (i) *in vivo*, from the caecum of mono-associated mice; (ii) *in vitro*, from bacteria grown in culture media; (iii) *in vitro*, from bacteria grown in the presence of dietary components; and (iv) from bacteria incubated on the surface of confluent Caco-2 and HT-29 cells.

Table S2. Real-time PCR analysis on bacterial RNA using *R. hominis* specific primers. Bacterial PCR primers were designed for differentially expressed genes using Primer3Plus. All samples were run in triplicate. *GyrA* was used as a reference gene for normalization.

Gene name	Rh14d - Control		Rh28d - Control		Rh14d - In Vitro+diet		Rh28d - In Vitro+diet		In Vitro + diet - In Vitro	
	FC	P-value	FC	P-value	FC	P-value	FC	P-value	FC	P-value
3-hydroxyacyl-CoA dehydrogenase	5.62	0.00393	8.62	0.00000	-0.85	0.69927	1.30	0.38670	6.62	0.00124
Acetyl-CoA acetyltransferase	10.25	0.00304	18.27	0.00000	1.38	0.51556	2.46	0.01882	7.44	0.00048
Aldose epimerase family protein	7.20	0.00002	11.09	0.00001	22.65	0.03173	34.89	0.02105	-0.32	0.30489
ATP synthase alpha chain	1.62	0.12713	2.50	0.00003	3.25	0.00645	5.00	0.00000	-2.00	0.00147
ATP synthase alpha chain2	11.94	0.00122	5.50	0.00001	2.71	0.18580	1.25	0.72876	4.41	0.06479
ATP synthase beta chain	1.85	0.08341	3.14	0.00056	4.00	0.00397	6.80	0.00001	-2.16	0.00241
ATP synthase beta chain2	8.27	0.00058	6.13	0.00025	1.64	0.33264	1.21	0.67923	-5.05	0.01400
ATP synthase gamma chain	2.31	0.02994	3.58	0.00027	4.22	0.00238	6.52	0.00001	-1.82	0.01276
ATP synthase gamma chain2	-9.59	0.00081	-10.55	0.00013	-1.59	0.40064	-1.75	0.29398	-6.02	0.01363
Butyryl-CoA dehydrogenase	13.31	0.00166	19.32	0.00000	1.48	0.39908	2.14	0.02840	9.03	0.00039
Electron transfer flavoprotein, alpha subunit	6.32	0.00729	15.55	0.00000	-0.80	0.61916	1.96	0.03441	7.94	0.00030
Electron transfer flavoprotein, beta subunit	7.94	0.00414	12.31	0.00000	1.13	0.78504	1.76	0.08347	7.00	0.00078
Flagellar motor rotation protein MotA	-0.94	0.71937	2.37	0.00023	1.69	0.02442	4.25	0.00001	-0.56	1.00000
Flagellar motor rotation protein MotB	-0.57	0.01030	1.19	0.00938	-0.75	0.14601	1.58	0.01676	-0.76	0.08011
Flagellin protein FlaA1	-3.05	0.00125	-1.93	0.01193	3.13	0.01261	4.95	0.00249	-9.57	0.00079
Flagellin protein FlaA2	1.04	0.84732	-1.28	0.05814	1.45	0.29909	1.09	0.77089	-1.39	0.28105
Flagellin protein FlaA3	1.14	0.41467	1.94	0.00312	1.63	0.08606	2.77	0.00477	-1.43	0.17174
Flagellin protein flaB	1.02	0.97319	-4.99	0.00568	1.96	0.40946	-2.60	0.14359	-1.92	0.25258
Glucuronide permease	-9.34	0.00001	-13.81	0.00022	-13.97	0.00001	-20.65	0.00003	-1.50	0.13959
L-threonine 3-O-phosphate decarboxylase	1.62	0.00963	3.70	0.00002	1.66	0.07350	3.78	0.00186	-0.98	0.92518
Magnesium transporter	372.00	0.00123	11.20	0.03391	4.42	0.32048	-0.13	0.18758	84.10	0.01672
Methyl-accepting chemotaxis protein1	-1.18	0.54522	-1.83	0.02731	-2.90	0.01211	-1.87	0.05778	-0.29	0.00389

Methyl-accepting protein2	chemotaxis	-2.46	0.00400	-2.95	0.00154	1.48	0.10070	1.24	0.33586	-3.65	0.00043
Methyl-accepting protein3	chemotaxis	1.13	0.54189	1.86	0.04504	3.17	0.00033	5.25	0.00055	-2.81	0.00020
Methyl-accepting sensory transducer1	chemotaxis	1.33	0.01743	2.15	0.00001	1.58	0.05704	2.56	0.00396	-1.19	0.38997
Methyl-accepting sensory transducer2	chemotaxis	2.03	0.00671	2.34	0.00017	4.94	0.00007	5.71	0.00003	-2.44	0.00163
MobA/MobL protein4/putative transfer protein	family conjugal	84.83	0.00177	5.77	0.01561	3.67	0.34227	-4.00	0.29305	23.10	0.04583
MobA/MobL protein1		257.93	0.00472	9.27	0.04437	8.72	0.25455	-3.19	0.50297	29.57	0.07733
MobA/MobL protein2		714.11	0.00172	11.64	0.08286	8.17	0.29225	-7.51	0.31871	87.43	0.04300
MobA/MobL protein3		362.26	0.00144	11.10	0.03880	7.77	0.27426	-4.20	0.43208	46.62	0.06260
MobA/MobL protein4		219.75	0.00147	6.52	0.08511	7.99	0.17452	-4.22	0.34388	27.49	0.04519
Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein oppA		1.26	0.49437	1.11	0.45125	1.28	0.49171	1.13	0.53525	-1.02	0.92102
Oligopeptide transport ATP-binding protein oppD		-1.29	0.05256	1.08	0.40875	-1.40	0.07422	-1.00	0.98737	1.09	0.60319
Osmosensitive K+ channel histidine kinase KdpD		3.98	0.00004	7.07	0.00000	5.58	0.00005	9.91	0.00001	-0.71	0.12049
Phosphate regulon sensor protein PhoR		2.73	0.00211	1.54	0.04166	3.80	0.00919	2.14	0.07191	-1.39	0.34941
Phosphoenolpyruvate carboxykinase [ATP]		2.30	0.05000	3.22	0.00004	1.15	0.70105	1.61	0.04623	2.00	0.00988
Potassium uptake protein, integral membrane component, KtrB		7.74	0.00006	10.81	0.00002	50.92	0.02191	71.15	0.01638	-0.15	0.15959
Putative conjugal transfer protein MobA/MobL		183.02	0.00023	9.49	0.01454	8.09	0.08695	-2.39	0.44355	22.63	0.02615
Pyruvate-flavodoxin oxidoreductase		-0.52	0.05848	-0.92	0.52386	-0.36	0.01504	-0.65	0.09510	1.42	0.16352
RNA polymerase sigma factor for flagellar operon		-0.76	0.02253	1.93	0.00032	-0.82	0.08962	2.08	0.00013	-0.93	0.40060

Table S3. Index of *R. hominis* PCR experiments plotted on circular genome map. List of *R. hominis* PCR experiments as shown on the circular genome map.

Exp	Primer	Start	End	Length	Gene
1	gyrA-N-F	7803	7889	87 bp	gyrA
2	1602-F	95664	95740	77 bp	Phosphate regulon sensor protein PhoR
3	1653-F	153403	153483	81 bp	Flagellin protein FlaA1
4	1686-F	189295	189382	88 bp	Flagellin protein FlaA2
5	1718-F	221205	221279	75 bp	Flagellin protein FlaA3
6	1735-F	250582	250674	93 bp	Methyl-accepting chemotaxis protein1
7	1769-F	290546	290628	83 bp	Methyl-accepting chemotaxis sensory transducer1
8	1770-F	291722	291808	87 bp	Methyl-accepting chemotaxis protein2
9	1831-N-F	348711	348810	100 bp	MobA/MobL protein4
10	1842-F	364775	364851	77 bp	MobA/MobL protein2
11	1867-2652-F	391044	391120	77 bp	MobA/MobL family protein4/putative conjugal transfer protein
12	2055-F	600837	600928	92 bp	Acetyl-CoA acetyltransferase
13	2056-F	602279	602363	85 bp	3-hydroxyacyl-CoA dehydrogenase
14	2057-F	602961	603037	77 bp	Butyryl-CoA dehydrogenase
15	2058-F	604411	604504	94 bp	Electron transfer flavoprotein, beta subunit
16	2059-F	605434	605516	83 bp	Electron transfer flavoprotein, alpha subunit
17	129-F	653987	654066	80 bp	Oligopeptide ABC transporter, periplasmic oligopeptide-binding protein oppA
18	132-F	658435	658516	82 bp	Oligopeptide transport ATP-binding protein oppD
19	805-R	934310	934406	97 bp	Osmosensitive K ⁺ channel histidine kinase KdpD
20	807-R	935306	935394	89 bp	Aldose epimerase family protein
21	808-R	936111	936190	80 bp	Potassium uptake protein, integral membrane component, KtrB
22	909-F	1053529	1053604	76 bp	Pyruvate-flavodoxin oxidoreductase
23	1235-F	1434705	1434785	81 bp	MobA/MobL protein3
24	1296-F	1495460	1495544	85 bp	Methyl-accepting chemotaxis protein3
25	1297-R	1497854	1497931	78 bp	L-threonine 3-O-phosphate decarboxylase
26	1335-F	1540579	1540671	93 bp	Flagellar motor rotation protein MotA
27	1336-F	1541416	1541511	96 bp	Flagellar motor rotation protein MotB
28	1356-F	1559143	1559227	85 bp	RNA polymerase sigma factor for flagellar operon
29	3119-F	2211612	2211705	94 bp	Phosphoenolpyruvate carboxykinase [ATP]
30	3117-R	2213046	2213139	94 bp	Methyl-accepting chemotaxis sensory transducer2
31	1867-2652-R	2736100	2736176	77 bp	MobA/MobL family protein4/putative conjugal transfer protein
32	1552M-R	2984489	2984566	78 bp	Magnesium transporter
33	397-R	3153341	3153427	87 bp	ATP synthase beta chain2
34	398-R	3153616	3153699	84 bp	ATP synthase gamma chain2
35	399-R	3155799	3155898	100 bp	ATP synthase alpha chain2
36	404-R	3159387	3159467	81 bp	Glucuronide permease
37	2399-R	3308172	3308265	94 bp	Putative conjugal transfer protein MobA/MobL
38	2323-R	3366526	3366615	90 bp	Flagellin protein flaB
39	2281-R	3416947	3417042	96 bp	ATP synthase beta chain
40	2280-R	3418736	3418824	89 bp	ATP synthase gamma chain
41	2279-R	3418857	3418942	86 bp	ATP synthase alpha chain
42	641-R	3467164	3467255	92 bp	MobA/MobL protein1

Fifty differentially expressed genes were identified (*in vivo* vs. *in vitro*). The most surprising discovery was the high up-regulation *in vivo* of genes involved in conjugation/mobilization transfer, the *mobA*- and *mobL*-like genes (Fig. 2A). The presence of such genes in the transcriptional studies was surprising as no identifiable genes were assigned to Phages, Prophages, Transposable Elements and Plasmids in the Subsystem Category feature. This difference in gene detection and allocation is likely due to the recognized limitations of the Subsystem Category annotation. The stimulatory effect of dietary compounds was much less pronounced, suggesting that the gut environment *per se* is a major inducer of genes involved in horizontal gene transfer.

Other gut environment-induced subsystems included Membrane Transport, in particular magnesium transport, and Motility and Chemotaxis including multiple methyl-accepting chemotaxis proteins and genes of the flagellar operon (Fig. 2B). *R. hominis* possesses multiple flagellin genes *flaA1*, *flaA2*, *flaA3*, and *flaB*. In the mouse gut environment, flagellin expression was verified by Western-blot and immunocytochemistry using bacteria isolated from both *in vivo* colonized mice and from *in vitro* cultures grown in the presence of diet, with specific antibodies raised against recombinant flagellin proteins FlaA1 and FlaA2 (RH1 and RH2) (Fig. 2C). This positive validation of flagellin expression *in vivo* is consistent with previous reports indicating that only certain subsets of Firmicutes produce flagella *in vivo* (Turnbaugh *et al.* 2008) unlike other bacterial species that actively down-regulate expression of the bacterial protein. The expression of catabolic metabolism genes in *R. hominis* in the gut environment was also affected by the gut environment (Fig. 2D). The genes involved included acetyl-CoA acetyltransferase, 3-hydroxyacyl-CoA dehydrogenase, butyryl-CoA dehydrogenase and phosphoenolpyruvate carboxykinase [ATP].

To further investigate the effects of host-adaptation on the *R. hominis* transcriptome, *in vitro* stimulation of human intestinal epithelial cells (Caco-2 and HT-29) was performed. This showed that the conjugation/mobilization transfer gene *mobA/mobL protein1*, which was induced by adaptation to the mouse gut, was also increased in both cell lines (Fig. 2E). Consistent with the *in vivo* data, the flagellin gene *MotA* was up-regulated in Caco-2 cells. Genes involved in butyrate metabolism showed differences between the two cell lines, with down-regulation seen in Caco-2 cells and up-regulation in HT-29 cells.

***R. hominis* influences host innate signalling pathways mostly in the colon**

Colonization of GF mice with *R. hominis* correlated with increased host gene expression, which was highest in the colon (Fig. 3A and Table S4). Differential expression was most

profound at 28d after colonization, with 159 genes up-regulated and 143 genes down-regulated. The number of differentially expressed genes in the ileum at 14d was similar to the ascending colon, with 79 genes up-regulated and 119 genes down-regulated. Differential expression in the ileum was very low at 28d, consistent with reduced colonization levels. The transcriptomic response differed at the two time-points, as shown by the clear separation of significant transcripts by heatmap analysis (Fig. 3B). Positive Real-time PCR validation of Affymetrix data is shown in Fig. 3C.

Table S4. Affymetrix data between *R. hominis* inoculated animals and GF animals.

Affymetrix microarray analysis was performed on RNA isolated from ascending colon and ileum tissue. Data was considered significant when $P < 0.05$ using the Benjamini and Hochberg false discovery method. (A) Transcripts differentially expressed in the ascending colon at 14d. (B) Transcripts differentially expressed in the ascending colon at 28d. (C) Transcripts differentially expressed in the ileum at 14d. (D) Transcripts differentially expressed in the ileum at 28d.

Table S4A. Transcripts differentially expressed in the ascending colon at 14d between *R. hominis* inoculated animals and germfree animals.

ID	Symbol	Name	FC	P-value	B
1424973_at	Cyp3a25	Cytochrome P450, family 3, subfamily a, polypeptide 25	26.35	4.09E-05	8.91
1449375_at	Ces6	Carboxylesterase 6	22.80	5.71E-16	29.28
1428400_at	2200002K05Rik	RIKEN cDNA 2200002K05 gene	6.22	6.65E-09	17.20
1419393_at	Abcg5	ATP-binding cassette, sub-family G (WHITE), member 5	5.81	0.002156	4.51
1429726_at	Slc16a9	Solute carrier family 16 (monocarboxylic acid transporters), member 9	4.38	0.001103	5.26
1430641_at	9030605I04Rik	RIKEN cDNA 9030605I04 gene	4.36	0.000358	6.52
1436575_at	Grin3a	Glutamate receptor ionotropic, NMDA3A	3.91	0.018503	1.71
1422749_at	Ly6g6c	Lymphocyte antigen 6 complex, locus G6C	3.71	0.004044	3.79
NuGO_emi070648_at	Abca12	ATP-binding cassette, sub-family A (ABC1), member 12	3.69	0.000333	6.62
1428682_at	Zc3h6	Zinc finger CCH type containing 6	3.28	0.002747	4.26
1431554_a_at	Anxa9	Annexin A9	3.00	3.25E-05	9.18
1423556_at	Akr1b7	Aldo-keto reductase family 1, member B7	2.87	0.018503	1.70
1424451_at	Acaa1b	Acetyl-Coenzyme A acyltransferase 1B	2.70	0.000404	6.33
1418486_at	Vnn1	Vanin 1	2.66	0.001531	4.93
1418606_at	Hoxd10	Homeo box D10	2.60	0.000102	7.97
1435207_at	Dixdc1	DIX domain containing 1	2.55	0.014337	2.11
1427072_at	Stard8	START domain containing 8	2.48	0.027058	1.15
1442560_at	NA	NA	2.42	0.006928	3.17
1420998_at	Etv5	Ets variant gene 5	2.33	0.017697	1.81
1440925_at	Rhoq	Ras homolog gene family, member Q	2.32	0.013564	2.22
1428902_at	Chst11	Carbohydrate sulfotransferase 11	2.30	0.000154	7.46
1416607_at	4931406C07Rik	RIKEN cDNA 4931406C07 gene	2.28	7.26E-11	21.08
1435673_at	Hoxd3	Homeo box D3	2.26	0.019944	1.60
1426663_s_at	Slc45a3	Solute carrier family 45, member 3	2.21	0.007554	3.07
1419651_at	2610200G18Rik	RIKEN cDNA 2610200G18 gene	2.21	0.000406	6.29
1422188_s_at	NA	NA	2.20	0.000397	6.39
1435468_at	D230025D16Rik	RIKEN cDNA D230025D16 gene	2.15	0.018503	1.71
1417991_at	Dio1	Deiodinase, iodothyronine, type I	2.12	0.026066	1.24
1451557_at	Tat	Tyrosine aminotransferase	2.11	0.046376	0.38

1428989_at	0710001D07Rik	RIKEN cDNA 0710001D07 gene	2.03	0.04028	0.54
1456680_at	B3gnt6	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 6 (core 3 synthase)	2.01	0.024447	1.35
1443235_at	Elf2ak4	Eukaryotic translation initiation factor 2 alpha kinase 4	1.97	0.002156	4.51
1419759_at	Abcb1a	ATP-binding cassette, sub-family B (MDR/TAP), member 1A	1.97	0.009639	2.79
1456389_at	Zeb2	Zinc finger E-box binding homeobox 2	1.96	0.013795	2.16
1440840_at	D630004K10Rik	RIKEN cDNA D630004K10 gene	1.96	0.017697	1.80
1457619_at	Ces6	Carboxylesterase 6	1.94	0.04028	0.55
1449049_at	Tlr1	Toll-like receptor 1	1.92	0.018503	1.72
NuGO_empt049113_at	Ptprh	Protein tyrosine phosphatase, receptor type, H	1.89	0.002961	4.17
1424376_at	Cdc42ep1	CDC42 effector protein (Rho GTPase binding) 1	1.88	0.014337	2.10
1436566_at	Rab40b	Rab40b, member RAS oncogene family	1.87	0.028805	1.04
1432590_at	LOC621549	NA	1.83	0.012459	2.38
1430543_at	Clip3	CAP-GLY domain containing linker protein 3	1.82	6.24E-05	8.47
1435553_at	Pdzd2	PDZ domain containing 2	1.81	0.048735	0.32
1428271_at	Acbd4	Acyl-Coenzyme A binding domain containing 4	1.80	0.014378	2.09
1418059_at	Eltld1	EGF, latrophilin seven transmembrane domain containing 1	1.79	0.005331	3.47
1456532_at	Pdgfd	Platelet-derived growth factor, D polypeptide	1.77	0.007554	3.05
1437393_at	A1875142	Expressed sequence A1875142	1.77	0.03504	0.78
1428332_at	1500004A08Rik	RIKEN cDNA 1500004A08 gene	1.77	0.038684	0.61
1434015_at	Slc2a6	Solute carrier family 2 (facilitated glucose transporter), member 6	1.75	0.000437	6.19
1449403_at	Pde9a	Phosphodiesterase 9A	1.71	0.03504	0.79
1428260_at	Spg3a	Spastic paraplegia 3A homolog (human)	1.69	0.011437	2.55
1417803_at	1110032A04Rik	RIKEN cDNA 1110032A04 gene	1.66	0.001103	5.27
1430245_at	Fxr1h	Fragile X mental retardation gene 1, autosomal homolog	1.66	0.012459	2.36
1422542_at	Gpr34	G protein-coupled receptor 34	1.65	0.031372	0.95
1427020_at	Scara3	Scavenger receptor class A, member 3	1.60	0.010425	2.69
1430211_at	4930415O20Rik	RIKEN cDNA 4930415O20 gene	1.57	0.045	0.42
1452809_at	9030607L17Rik	RIKEN cDNA 9030607L17 gene	1.53	0.013564	2.22
NuGO_empt066282_at	Defb37	Defensin beta 37	1.51	0.005422	3.44
1451498_at	BC004853	cDNA sequence BC004853	1.50	0.037045	0.69
1434140_at	Mcf2l	Mcf.2 transforming sequence-like	1.49	0.025597	1.30
1452650_at	Trim62	Tripartite motif-containing 62	1.46	0.02652	1.19
1427492_at	Pof1b	Premature ovarian failure 1B	1.46	0.004679	3.61

1426601_at	Slc37a1	Solute carrier family 37 (glycerol-3-phosphate transporter), member 1	1.42	0.032433	0.90
1448188_at	Ucp2	Uncoupling protein 2 (mitochondrial, proton carrier)	1.42	0.008209	2.96
1460409_at	Cpt1a	Carnitine palmitoyltransferase 1a, liver	1.40	0.025932	1.26
1460652_at	Esrra	Estrogen related receptor, alpha	1.39	0.038182	0.65
1453869_at	LOC328277	NA	1.36	0.017271	1.86
1435985_at	Farp2	FERM, RhoGEF and pleckstrin domain protein 2	1.34	0.037045	0.70
1454706_at	Uvrag	UV radiation resistance associated gene	1.34	0.021704	1.51
1451232_at	Cd151	CD151 antigen	1.31	0.049274	0.31
1423570_at	Abcg1	ATP-binding cassette, sub-family G (WHITE), member 1	1.30	0.04028	0.55
1418817_at	Chmp1b	Chromatin modifying protein 1B	1.25	0.010981	2.61
1441059_at	1700049G17Rik	RIKEN cDNA 1700049G17 gene	1.25	0.028805	1.05
NuGO_empt061346_at	NA	NA	1.24	0.026697	1.17
1420503_at	Slc6a14	Solute carrier family 6 (neurotransmitter transporter), member 14	1.22	0.012499	2.32
NuGO_empt073103_x_at	NA	NA	1.21	0.01946	1.63
1452592_at	Mgst2	Microsomal glutathione S-transferase 2	1.18	0.038684	0.60
1441135_at	NA	NA	-1.10	0.01646	1.92
1428482_at	Akap10	A kinase (PRKA) anchor protein 10	-1.20	0.022427	1.47
1426679_at	Zfp706	Zinc finger protein 706	-1.26	0.028805	1.05
1415816_at	Cct7	Chaperonin subunit 7 (eta)	-1.26	0.011936	2.47
1445024_at	Stard7	START domain containing 7	-1.27	0.03504	0.79
1453591_at	5730437N04Rik	RIKEN cDNA 5730437N04 gene	-1.32	0.02652	1.19
1450788_at	Saa1	Serum amyloid A 1	-1.32	0.002156	4.54
1436157_at	Ccar1	Cell division cycle and apoptosis regulator 1	-1.33	0.034783	0.81
1450987_a_at	2310004I24Rik	RIKEN cDNA 2310004I24 gene	-1.35	0.0361	0.73
1433850_at	Ppp4r2	Protein phosphatase 4, regulatory subunit 2	-1.36	0.018503	1.71
1434835_at	Wapal	Wings apart-like homolog (Drosophila)	-1.37	0.027058	1.14
NuGO_empt090017_s_at	Cdk7	Cyclin-dependent kinase 7 (homolog of Xenopus MO15 cdk-activating kinase)	-1.37	0.044186	0.44
1422579_at	Hspe1	Heat shock protein 1 (chaperonin 10)	-1.38	0.017697	1.81
1455726_at	Gm71	Gene model 71, (NCBI)	-1.42	0.029302	1.02
1420461_at	Mst1r	Macrophage stimulating 1 receptor (c-met-related tyrosine kinase)	-1.49	0.046838	0.37
1453264_at	Marveld3	MARVEL (membrane-associating) domain containing 3	-1.49	0.005522	3.40

1425030_at	Zfp622	Zinc finger protein 622	-1.49	0.000586	5.89
1458507_at	2810055G22Rik	RIKEN cDNA 2810055G22 gene	-1.50	0.032433	0.89
1443988_at	Rbm39	RNA binding motif protein 39	-1.51	0.037045	0.68
1460465_at	A930038C07Rik	RIKEN cDNA A930038C07 gene	-1.51	0.004627	3.64
1428784_at	Gmip	Gem-interacting protein	-1.51	0.045	0.42
1451621_at	5830417C01Rik	RIKEN cDNA 5830417C01 gene	-1.65	0.014536	2.07
1422837_at	Scel	Sciellin	-1.68	0.0361	0.74
1454617_at	Arddc3	Arrestin domain containing 3	-1.71	0.038496	0.62
1452047_at	Cacybp	Calcyclin binding protein	-1.72	0.018495	1.75
1434724_at	Usp31	Ubiquitin specific peptidase 31	-1.75	0.024018	1.39
1417032_at	Ube2g2	Ubiquitin-conjugating enzyme E2G 2	-1.76	0.04028	0.56
1448628_at	Scg3	Secretogranin III	-1.77	0.016351	1.93
1443877_a_at	Rapgef6	Rap guanine nucleotide exchange factor (GEF) 6	-1.82	0.004044	3.80
1427944_at	Caprin2	Caprin family member 2	-1.85	0.04248	0.49
1415909_at	Stip1	Stress-induced phosphoprotein 1	-1.89	0.024018	1.38
1422452_at	Bag3	Bcl2-associated athanogene 3	-1.89	0.026066	1.24
1438041_at	Pde7a	Phosphodiesterase 7A	-1.91	0.025597	1.29
1433927_at	Usp11	Ubiquitin specific peptidase like 1	-1.92	0.024018	1.38
1422860_at	Nlts	Neurotensin	-1.94	1.84E-07	14.19
1451194_at	Aldob	Aldolase 2, B isoform	-1.94	7.08E-06	10.59
1441662_at	Cyp4x1	Cytochrome P450, family 4, subfamily x, polypeptide 1	-1.95	0.00019	7.22
1445490_at	C77805	Expressed sequence C77805	-1.96	0.003479	4.00
NuGO_emt066852_at	NA	NA	-2.07	0.013658	2.19
1442427_at	NA	NA	-2.11	0.033207	0.86
1430185_at	Akap13	A kinase (PRKA) anchor protein 13	-2.12	0.0361	0.74
1452462_a_at	Banp	Btg3 associated nuclear protein	-2.16	0.038182	0.64
1446158_at	Exoc6b	Exocyst complex component 6B	-2.19	0.010302	2.72
1418113_at	Cyp2d10	Cytochrome P450, family 2, subfamily d, polypeptide 10	-2.24	0.012383	2.42
1426645_at	Hsp90aa1	Heat shock protein 90kDa alpha (cytosolic), class A member 1	-2.30	0.012459	2.34
1420150_at	Spsb1	SplA/ryanodine receptor domain and SOCS box containing 1	-2.36	0.015257	2.01
1452382_at	Dnm3os	Dynamitin 3, opposite strand	-2.37	0.025932	1.26
1460645_at	Chordc1	Cysteine and histidine-rich domain (CHORD)-containing, zinc-binding protein 1	-2.38	0.002156	4.53
1457477_at	NA	NA	-2.39	0.027058	1.14

1430175_at	Tmtc2	Transmembrane and tetra-trico-peptide repeat containing 2	-2.40	0.028805	1.05
1433266_at	2810416A17Rik	RIKEN cDNA 2810416A17 gene	-2.43	0.0361	0.73
1416756_at	Dnajb1	DnaJ (Hsp40) homolog, subfamily B, member 1	-2.50	0.013658	2.19
1415938_at	Spink3	Serine peptidase inhibitor, Kazal type 3	-2.53	0.012459	2.35
1429273_at	Bmper	BMP-binding endothelial regulator	-2.53	0.012459	2.34
1450518_at	Hnf4g	Hepatocyte nuclear factor 4, gamma	-2.53	0.019018	1.66
1440227_at	BF642829	Expressed sequence BF642829	-2.58	0.038494	0.63
1451924_a_at	Edn1	Endothelin 1	-2.63	0.011568	2.51
1425952_a_at	Gcg	Glucagon	-2.65	2.76E-06	11.57
1459253_at	Arrdc3	Arrestin domain containing 3	-2.69	0.026081	1.23
1416288_at	Dnaja1	DnaJ (Hsp40) homolog, subfamily A, member 1	-2.75	0.000154	7.45
1435160_at	1110064P04Rik	RIKEN cDNA 1110064P04 gene	-2.76	0.011568	2.52
1460179_at	Dnaja1	DnaJ (Hsp40) homolog, subfamily A, member 1	-2.78	0.000251	6.92
1419349_a_at	Cyp2d9	Cytochrome P450, family 2, subfamily d, polypeptide 9	-2.78	0.000134	7.68
NuGO_empt044299_at	Sstr1	Somatostatin receptor 1	-2.80	3.73E-07	13.38
1449493_at	InsI5	Insulin-like 5	-2.91	5.33E-08	15.35
NuGO_empt054105_at	D630013G24Rik	RIKEN cDNA D630013G24 gene	-2.99	0.004044	3.83
1419185_a_at	Mlxip1	MLX interacting protein-like	-3.16	4.43E-06	11.08
1449939_s_at	Dlk1	Delta-like 1 homolog (Drosophila)	-3.29	9.19E-11	20.69
1458385_at	Hspa4l	Heat shock protein 4 like	-3.36	0.032433	0.90
1422639_at	Calcb	Calcitonin-related polypeptide, beta	-3.75	0.010981	2.62
1425993_a_at	Hsp110	Heat shock protein 110	-4.08	0.028805	1.05
NuGO_empt091063_at	NA	NA	-6.03	3.11E-07	13.63
1419473_a_at	Cck	Cholecystokinin	-10.87	4.94E-11	21.66
1452388_at	Hspa1a	Heat shock protein 1A	-25.30	0.02652	1.19

Table S4B. Transcripts differentially expressed in the ascending colon at 28d between *R. hominis* inoculated animals and germfree animals.

ID	Symbol	Name	FC	P-value	B
1416809_at	Cyp3a11	Cytochrome P450, family 3, subfamily a, polypeptide 11	89.84	1.45E-06	12.00
1424973_at	Cyp3a25	Cytochrome P450, family 3, subfamily a, polypeptide 25	52.30	7.31E-07	12.67
1419393_at	Abcg5	ATP-binding cassette, sub-family G (WHITE), member 5	42.06	1.60E-09	18.55
1449375_at	Ces6	Carboxylesterase 6	26.10	1.75E-16	31.66
1422749_at	Ly6g6c	Lymphocyte antigen 6 complex, locus G6C	25.49	2.06E-10	20.64

1448964_at	S100g	S100 calcium binding protein G	17.88	4.82E-07	13.11
1455540_at	Cps1	Carbamoyl-phosphate synthetase 1	11.50	0.010507	1.76
1449133_at	Sprr1a	Small proline-rich protein 1A	11.27	0.009303	1.97
NuGO_emt070648_at	Abca12	ATP-binding cassette, sub-family A (ABC1), member 12	7.97	4.05E-08	15.43
1448485_at	Ggt1	Gamma-glutamyltransferase 1	7.45	0.020934	0.94
1417828_at	Aqp8	Aquaporin 8	6.67	0.001582	4.30
1437060_at	Olfm4	Olfactomedin 4	5.80	0.023469	0.77
1420437_at	Indo	Indoleamine-pyrrole 2,3 dioxygenase	5.18	0.000637	5.44
1425452_s_at	AW125753	Expressed sequence AW125753	4.97	0.001639	4.20
1439934_at	Slc30a10	Solute carrier family 30, member 10	4.68	0.000119	7.49
1430641_at	9030605I04Rik	RIKEN cDNA 9030605I04 gene	4.47	0.000199	6.89
1423556_at	Akr1b7	Aldo-keto reductase family 1, member B7	4.36	0.000304	6.44
1428400_at	2200002K05Rik	RIKEN cDNA 2200002K05 gene	4.31	4.81E-07	13.18
1424626_at	2010003K11Rik	RIKEN cDNA 2010003K11 gene	4.25	0.024887	0.68
1439727_at	Cla6	Chloride channel calcium activated 6	4.09	0.000869	5.07
1450355_a_at	Capg	Capping protein (actin filament), gelsolin-like	3.90	4.67E-05	8.59
1427119_at	Spink4	Serine peptidase inhibitor, Kazal type 4	3.67	0.007507	2.25
NuGO_emt092118_s_at	NA	NA	3.51	0.040399	-0.04
1451239_a_at	Slc26a1	Solute carrier family 26 (sulfate transporter), member 1	3.36	0.004508	2.94
1418283_at	Cldn4	Claudin 4	2.97	0.002809	3.51
1418165_at	Itlna	Intelectin a	2.88	0.0005	5.82
1440192_at	1810054D07Rik	RIKEN cDNA 1810054D07 gene	2.86	0.000108	7.65
1426980_s_at	E130012A19Rik	RIKEN cDNA E130012A19 gene	2.83	0.002052	3.91
1431042_at	Pagr8	Progestin and adipoQ receptor family member VIII	2.67	0.000637	5.46
1424688_at	Creb3l3	cAMP responsive element binding protein 3-like 3	2.67	0.035165	0.19
NuGO_emt049113_at	Ptprh	Protein tyrosine phosphatase, receptor type, H	2.54	6.75E-06	10.55
1432358_at	Muc16	Mucin 16	2.47	0.044391	-0.20
1419759_at	Abcb1a	ATP-binding cassette, sub-family B (MDR/TAP), member 1A	2.46	0.00017	7.13
NuGO_emt033610_at	Nox1	NADPH oxidase 1	2.43	0.035165	0.20
1418661_at	Abhd2	Abhydrolase domain containing 2	2.41	0.043727	-0.17
1420499_at	Gch1	GTP cyclohydrolase 1	2.38	0.040391	-0.03
1416607_at	4931406C07Rik	RIKEN cDNA 4931406C07 gene	2.31	3.39E-11	22.26
1417991_at	Dio1	Deiodinase, iodothyronine, type I	2.29	0.006845	2.39
1455455_at	Glit2d2	Glycosyltransferase 28 domain containing 2	2.16	0.015457	1.33

1417164_at	Dusp10	Dual specificity phosphatase 10	2.13	0.031006	0.40
1428937_at	Atp2b1	ATPase, Ca++ transporting, plasma membrane 1	2.13	1.13E-05	9.98
1429833_at	Ly6g6e	Lymphocyte antigen 6 complex, locus G6E	2.10	0.042694	-0.12
1419582_at	Cyp2c55	Cytochrome P450, family 2, subfamily c, polypeptide 55	2.06	0.022426	0.83
1448562_at	Upp1	Uridine phosphorylase 1	2.06	0.007369	2.29
1444254_at	NA	NA	2.05	0.009642	1.90
1428936_at	Atp2b1	ATPase, Ca++ transporting, plasma membrane 1	2.05	0.000378	6.15
1421268_at	Ugcg	UDP-glucose ceramide glucosyltransferase	2.05	0.035029	0.21
1419478_at	Sectm1b	Secreted and transmembrane 1B	2.02	0.004814	2.83
1428336_at	Agpat4	1-acylglycerol-3-phosphate O-acyltransferase 4 (lysophosphatidic acid acyltransferase, delta)	2.02	0.047797	-0.32
1456231_at	Pla2g3	Phospholipase A2, group III	1.99	0.000365	6.21
1421709_a_at	Fmo5	Flavin containing monooxygenase 5	1.97	0.02187	0.86
1455104_at	NA	NA	1.95	0.010934	1.71
1417133_at	Pmp22	Peripheral myelin protein	1.95	0.027135	0.58
1418206_at	Sdf2l1	Stromal cell-derived factor 2-like 1	1.94	0.045914	-0.25
1436614_at	Al843639	Expressed sequence Al843639	1.94	0.006795	2.41
1452070_at	Dedd2	Death effector domain-containing DNA binding protein 2	1.92	0.046848	-0.28
1417404_at	Elov16	ELOVL family member 6, elongation of long chain fatty acids (yeast)	1.90	0.012549	1.55
1417277_at	Cyp4f16	Cytochrome P450, family 4, subfamily f, polypeptide 16	1.87	0.000119	7.50
1422983_at	Itgb6	Integrin beta 6	1.87	0.03266	0.33
1454746_at	Plekhn1	Pleckstrin homology domain containing, family M (with RUN domain) member 1	1.86	0.001042	4.88
1425079_at	Tm6sf2	Transmembrane 6 superfamily member 2	1.85	0.001269	4.60
1455099_at	Mogat2	Monoacylglycerol O-acyltransferase 2	1.85	0.024887	0.68
1435749_at	Gda	Guanine deaminase	1.82	0.002467	3.67
1416488_at	Ccng2	Cyclin G2	1.81	0.003976	3.09
1418256_at	Srf	Serum response factor	1.79	0.033108	0.26
1426744_at	Srebf2	Sterol regulatory element binding factor 2	1.76	0.041026	-0.07
1457253_at	Trim40	Tripartite motif-containing 40	1.75	0.001269	4.61
1433556_at	Centa1	Centaurin, alpha 1	1.75	0.009141	2.01
NuGO_emt084792_x_at	NA	NA	1.74	0.020944	0.92
1417823_at	Gcat	Glycine C-acetyltransferase (2-amino-3-ketobutyrate-coenzyme A ligase)	1.74	0.049235	-0.38

NuGO_empt066282_at	Defb37	Defensin beta 37	1.74	6.42E-05	8.25
1429550_at	Entpd8	Ectonucleoside triphosphate diphosphohydrolase 8	1.73	0.004814	2.85
1430594_at	Rab11fip1	RAB11 family interacting protein 1 (class I)	1.73	0.033108	0.27
1420913_at	Slco2a1	Solute carrier organic anion transporter family, member 2a1	1.73	0.001791	4.09
NuGO_empt043440_at	2210010C17Rik	RIKEN cDNA 2210010C17 gene	1.73	0.01702	1.21
1430674_at	1700016C15Rik	RIKEN cDNA 1700016C15 gene	1.73	0.036022	0.14
1430890_at	2210010C17Rik	RIKEN cDNA 2210010C17 gene	1.72	0.011201	1.68
1417803_at	1110032A04Rik	RIKEN cDNA 1110032A04 gene	1.71	0.000358	6.25
1449873_at	Bmp8a	Bone morphogenetic protein 8a	1.71	0.044796	-0.22
1434130_at	Lhfp12	Lipoma HMGIC fusion partner-like 2	1.71	0.004814	2.83
1448605_at	Rhoc	Ras homolog gene family, member C	1.70	0.001099	4.79
1432363_at	2410018E23Rik	RIKEN cDNA 2410018E23 gene	1.69	0.04853	-0.34
1427878_at	0610010O12Rik	RIKEN cDNA 0610010O12 gene	1.68	2.48E-05	9.21
1416379_at	Panx1	Pannexin 1	1.67	0.002404	3.74
1434059_at	B230312A22Rik	RIKEN cDNA B230312A22 gene	1.65	0.001763	4.12
1452475_at	Pcsk5	Proprotein convertase subtilisin/kexin type 5	1.65	0.036022	0.12
1454399_at	2010003H20Rik	RIKEN cDNA 2010003H20 gene	1.62	0.016629	1.24
1460550_at	Mttr11	Myotubularin related protein 11	1.62	0.010762	1.73
NuGO_empt070892_at	NA	NA	1.62	0.020773	0.95
1436710_at	Zswim4	Zinc finger, SWIM domain containing 4	1.62	0.009718	1.88
1420663_at	Zbtb7b	Zinc finger and BTB domain containing 7B	1.61	0.042694	-0.13
1418991_at	Bak1	BCL2-antagonist/killer 1	1.61	0.01094	1.71
1417990_at	Ppp1r14d	Protein phosphatase 1, regulatory (inhibitor) subunit 14D	1.59	0.003746	3.16
1452837_at	Lpin2	Lipin 2	1.58	0.035702	0.16
NuGO_empt021769_s_at	Slc17a4	Solute carrier family 17 (sodium phosphate), member 4	1.57	0.015878	1.30
1418671_at	Capn5	Calpain 5	1.57	0.041629	-0.09
1417751_at	Stk10	Serine/threonine kinase 10	1.57	0.027636	0.55
1452294_at	Pcdh1	Protocadherin 1	1.56	0.003042	3.42
1429154_at	Slc35f2	Solute carrier family 35, member F2	1.56	0.004185	3.02
1450982_at	Slc9a3r1	Solute carrier family 9 (sodium/hydrogen exchanger), isoform 3 regulator 1	1.56	0.008755	2.06
1434015_at	Slc2a6	Solute carrier family 2 (facilitated glucose transporter), member 6	1.55	0.005093	2.72
1418712_at	Cdc42ep5	CDC42 effector protein (Rho GTPase binding) 5	1.55	0.003976	3.09
1424809_at	Crb3	Crumbs homolog 3 (Drosophila)	1.53	0.014107	1.43

1428953_at	Otud7b	OTU domain containing 7B	1.53	0.035702	0.17
1424090_at	Sdcbp2	Syndecan binding protein (syntenin) 2	1.53	0.023574	0.76
1418215_at	Mep1b	Meprin 1 beta	1.53	0.027636	0.55
1434456_at	Gm440	gene model 440, (NCBI)	1.53	0.032752	0.32
1423521_at	Lmnb1	Lamin B1	1.53	0.007369	2.31
1425298_a_at	Naip1	NLR family, apoptosis inhibitory protein 1	1.51	0.002438	3.70
1456619_at	Liph	Lipase, member H	1.50	0.044796	-0.22
1418976_s_at	Cideb	Cell death-inducing DNA fragmentation factor, alpha subunit-like effector B	1.49	0.020944	0.93
1423376_a_at	Dok4	Docking protein 4	1.49	0.005089	2.73
1415793_at	Pnpo	Pyridoxine 5'-phosphate oxidase	1.48	0.00252	3.63
1435461_at	Magi3	Membrane associated guanylate kinase, WW and PDZ domain containing 3	1.48	0.001623	4.23
1444951_at	BC042698	cDNA sequence BC042698	1.45	0.042694	-0.14
1452214_at	Skil	SKI-like	1.45	0.020383	0.99
1426284_at	Krt20	Keratin 20	1.43	0.021519	0.88
1460406_at	AI427122	Expressed sequence AI427122	1.42	0.016225	1.27
1419331_at	Cdh17	Cadherin 17	1.41	0.009399	1.94
1428509_at	Myo1e	Myosin IE	1.41	0.043727	-0.17
1429117_at	Tradd	TNFRSF1A-associated via death domain	1.41	0.020944	0.91
1460681_at	Ceacam2	CEA-related cell adhesion molecule 2	1.41	0.036022	0.14
1455678_at	NA	NA	1.41	0.048555	-0.36
1440218_at	BC040758	cDNA sequence BC040758	1.41	0.026321	0.62
1416009_at	Tspan3	Tetraspanin 3	1.40	0.002052	3.91
1456200_at	Ipmk	Inositol polyphosphate multikinase	1.40	0.033108	0.28
1424126_at	Alas1	Aminolevulinic acid synthase 1	1.39	0.040391	-0.03
1434482_at	D4Ertid22e	DNA segment, Chr 4, ERATO Doi 22, expressed	1.39	0.038213	0.05
1416690_at	Gtpbp2	GTP binding protein 2	1.38	0.035702	0.16
1417895_a_at	Tmem54	Transmembrane protein 54	1.38	0.001332	4.53
1424245_at	Ces2	Carboxylesterase 2	1.37	0.020383	0.99
1434559_at	Stx3	Syntaxin 3	1.37	0.032755	0.31
1426733_at	Itpk1	Inositol 1,3,4-triphosphate 5/6 kinase	1.35	0.010099	1.81
1451139_at	Slc39a4	Solute carrier family 39 (zinc transporter), member 4	1.34	0.04853	-0.34
1417398_at	Rras2	Related RAS viral (r-ras) oncogene homolog 2	1.34	0.040085	-0.02

1427203_at	Myo15b	Myosin XVB	1.33	0.020944	0.92
1428331_at	2210016F16Rik	RIKEN cDNA 2210016F16 gene	1.32	0.047659	-0.30
1427128_at	Ptpn23	Protein tyrosine phosphatase, non-receptor type 23	1.31	0.042694	-0.13
1420426_at	Myo7b	Myosin VIIb	1.30	0.036022	0.13
1452304_a_at	Arhgef5	Rho guanine nucleotide exchange factor (GEF) 5	1.30	0.047797	-0.31
1434303_at	Raph1	Ras association (RalGDS/AF-6) and pleckstrin homology domains 1	1.29	0.024887	0.69
1433885_at	AI788777	Expressed sequence AI788777	1.28	0.018321	1.10
1415765_at	Hnrpul2	Heterogeneous nuclear ribonucleoprotein U-like 2	1.28	0.009181	2.00
1448110_at	Sema4a	Sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain, (semaphorin) 4A	1.27	0.00151	4.40
1423960_at	Mboat5	Membrane bound O-acyltransferase domain containing 5	1.27	0.043727	-0.18
1415676_a_at	Psmb5	Proteasome (prosome, macropain) subunit, beta type 5	1.27	0.040399	-0.04
1434345_at	Clrn3	Clarin 3	1.27	0.027133	0.59
1426014_a_at	Mucdhl	Mucin and cadherin like	1.27	0.036152	0.11
1420503_at	Slc6a14	Solute carrier family 6 (neurotransmitter transporter), member 14	1.26	0.001582	4.31
1418817_at	Chmp1b	Chromatin modifying protein 1B	1.25	0.004961	2.78
1423686_a_at	Prr13	Proline rich 13	1.25	0.026919	0.60
1420826_at	Letm1	Leucine zipper-EF-hand containing transmembrane protein 1	1.24	0.036022	0.13
1448618_at	Mvp	Major vault protein	1.24	0.031006	0.39
1417178_at	Gipc2	GIPC PDZ domain containing family, member 2	1.24	0.014863	1.38
1416627_at	Spint1	Serine protease inhibitor, Kunitz type 1	1.23	0.048555	-0.36
1428163_at	Sar1b	SAR1 gene homolog B (S. cerevisiae)	1.22	0.036022	0.13
1416193_at	Car1	Carbonic anhydrase 1	1.21	0.009728	1.87
1444884_at	Ppt1	Palmitoyl-protein thioesterase 1	1.18	0.032755	0.31
1448279_at	Arpc3	Actin related protein 2/3 complex, subunit 3	1.14	0.049367	-0.38
1417282_at	Mmp23	Matrix metalloproteinase 23	-1.14	0.024366	0.72
1429615_at	Zfp91	Zinc finger protein 91	-1.22	0.038424	0.04
1453577_at	2610018I03Rik	RIKEN cDNA 2610018I03 gene	-1.22	0.04853	-0.35
1417999_at	Itim2b	Integral membrane protein 2B	-1.22	0.020944	0.91
1454994_at	Klhl20	Kelch-like 20 (Drosophila)	-1.24	0.040433	-0.05
1435563_at	Mrps5	Mitochondrial ribosomal protein S5	-1.25	0.022222	0.84
1445561_at	NA	NA	-1.27	0.04853	-0.34
1436854_at	Trpc2	Transient receptor potential cation channel, subfamily C, member 2	-1.27	0.002008	3.97
1416452_at	Oat	Ornithine aminotransferase	-1.27	0.009757	1.86

1415961_at	Itm2c	Integral membrane protein 2C	-1.28	0.007369	2.28
1448933_at	Pcdhb17	Protocadherin beta 17	-1.28	0.048555	-0.35
1440391_at	Gcn11	GCN1 general control of amino-acid synthesis 1-like 1 (yeast)	-1.29	0.030585	0.42
1450788_at	Saa1	Serum amyloid A 1	-1.31	0.001623	4.23
1444451_at	Pappa2	Pappalysin 2	-1.31	0.038548	0.03
1457029_at	C030010B13Rik	RIKEN cDNA C030010B13 gene	-1.31	0.023793	0.74
1417088_at	Zfp346	Zinc finger protein 346	-1.33	0.039077	0.01
1436220_at	Zfp287	Zinc finger protein 287	-1.33	0.022441	0.82
1445824_at	Zfp458	Zinc finger protein 458	-1.37	0.044011	-0.18
1420191_s_at	D16Bwg1494e	DNA segment, Chr 16, Brigham & Women's Genetics 1494 expressed	-1.37	0.035165	0.19
1415871_at	Tgfb1	Transforming growth factor, beta induced	-1.38	0.039653	0.00
1442731_at	9030416H16Rik	RIKEN cDNA 9030416H16 gene	-1.41	0.012107	1.59
1424889_at	Nupl2	Nucleoporin like 2	-1.42	0.001582	4.31
1416178_a_at	Plekhhb1	Pleckstrin homology domain containing, family B (evecitins) member 1	-1.44	0.047797	-0.32
1447946_at	Adam23	A disintegrin and metallopeptidase domain 23	-1.44	0.017961	1.13
NuGO_emt080869_at	NA	NA	-1.45	0.032755	0.31
1452050_at	Camk1d	Calcium/calmodulin-dependent protein kinase ID	-1.45	0.006002	2.56
1442447_at	Znrf3	Zinc and ring finger 3	-1.45	9.00E-06	10.24
1416865_at	Fgd1	FYVE, RhoGEF and PH domain containing 1	-1.46	0.049235	-0.37
1442197_at	Al480624	Expressed sequence Al480624	-1.47	0.033108	0.27
1427020_at	Scara3	Scavenger receptor class A, member 3	-1.47	0.029906	0.45
1434961_at	Asb1	Ankyrin repeat and SOCS box-containing protein 1	-1.48	0.0093	1.97
1431873_a_at	Tube1	Epsilon-tubulin 1	-1.48	0.033108	0.26
1424367_a_at	Homer2	Homer homolog 2 (Drosophila)	-1.49	0.00269	3.56
1441662_at	Cyp4x1	Cytochrome P450, family 4, subfamily x, polypeptide 1	-1.50	0.028455	0.51
1429086_at	Grhl2	Grainyhead-like 2 (Drosophila)	-1.51	0.04298	-0.15
1439078_at	Klhl4	Keich-like 4 (Drosophila)	-1.52	0.020679	0.96
1451194_at	Aldob	Aldolase 2, B isoform	-1.54	0.002438	3.70
1449913_at	Zfp2	Zinc finger protein 2	-1.55	0.033108	0.29
1431820_at	4632404H12Rik	RIKEN cDNA 4632404H12 gene	-1.56	0.036071	0.12
1437900_at	4930523C07Rik	RIKEN cDNA 4930523C07 gene	-1.56	0.044391	-0.20
1449462_at	3110049J23Rik	RIKEN cDNA 3110049J23 gene	-1.57	0.041709	-0.09
1457373_at	Cdh19	Cadherin 19, type 2	-1.57	0.032755	0.31
1423072_at	6720475J19Rik	RIKEN cDNA 6720475J19 gene	-1.58	0.006244	2.51

1422542_at	Gpr34	G protein-coupled receptor 34	-1.58	0.040399	-0.04
1448475_at	Olfml3	Olfactomedin-like 3	-1.58	0.032964	0.30
1417676_a_at	Ptpro	Protein tyrosine phosphatase, receptor type, O	-1.59	0.001623	4.24
1456763_at	AA536749	Expressed sequence AA536749	-1.59	0.017774	1.15
1417732_at	Anxa8	Annexin A8	-1.59	0.027267	0.57
1425510_at	Mark1	MAP/microtubule affinity-regulating kinase 1	-1.60	0.004814	2.85
1417234_at	Mmp11	Matrix metalloproteinase 11	-1.61	0.036152	0.11
1416194_at	Cyp4b1	Cytochrome P450, family 4, subfamily b, polypeptide 1	-1.62	0.002404	3.74
1429679_at	Fbxl13	F-box and leucine-rich repeat protein 13	-1.64	0.010507	1.76
1428260_at	Spg3a	Spastic paraplegia 3A homolog (human)	-1.68	0.007369	2.28
1426413_at	Neurod1	Neurogenic differentiation 1	-1.68	0.009718	1.88
1455500_at	Rnf213	Ring finger protein 213	-1.68	0.017961	1.13
1456532_at	Pdgfd	Platelet-derived growth factor, D polypeptide	-1.70	0.008375	2.11
1419754_at	Myo5a	Myosin Va	-1.71	0.020595	0.97
1460147_at	NA	NA	-1.71	0.037207	0.08
1440014_at	Pacs1	Phosphofurin acidic cluster sorting protein 1	-1.72	7.34E-05	8.05
1451342_at	Spon1	Spondin 1, (f-spondin) extracellular matrix protein	-1.73	0.045023	-0.23
1438530_at	Tfpi	Tissue factor pathway inhibitor	-1.76	0.001582	4.29
1449563_at	Cntn1	Contactin 1	-1.77	0.029208	0.48
1435829_at	B930008K04Rik	RIKEN cDNA B930008K04 gene	-1.78	0.004961	2.78
NuGO_emt010210_at	Cacna2d2	Calcium channel, voltage-dependent, alpha 2/delta subunit 2	-1.80	0.035702	0.17
1455633_at	Zfp647	Zinc finger protein 647	-1.80	0.041245	-0.08
1420416_at	Sema3a	Sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3A	-1.81	0.000506	5.78
1417644_at	Sspn	Sarcospan	-1.83	0.003984	3.08
1419687_at	D930010J01Rik	RIKEN cDNA D930010J01 gene	-1.83	0.045963	-0.25
1439618_at	Pde10a	Phosphodiesterase 10A	-1.83	0.017774	1.16
1440430_at	Elp4	Elongation protein 4 homolog (S. cerevisiae)	-1.84	0.023065	0.79
1425069_at	BC018285	cDNA sequence BC018285	-1.84	0.042264	-0.11
1419376_at	1110018M03Rik	RIKEN cDNA 1110018M03 gene	-1.85	0.009718	1.88
1434194_at	Mtap2	Microtubule-associated protein 2	-1.85	0.007048	2.36
1459707_at	Pacs1	Phosphofurin acidic cluster sorting protein 1	-1.86	0.004814	2.86
1434475_at	Ppig	Peptidyl-prolyl isomerase G (cyclophilin G)	-1.86	0.045576	-0.24
1449158_at	Kcnk2	Potassium channel, subfamily K, member 2	-1.87	0.004961	2.77

1460606_at	Hsd17b13	Hydroxysteroid (17-beta) dehydrogenase 13	-1.88	0.003153	3.38
1436087_at	Dpp10	Dipeptidylpeptidase 10	-1.89	0.043727	-0.17
NuGO_emt029633_at	Npy2r	Neuropeptide Y receptor Y2	-1.90	0.000199	6.88
1418606_at	Hoxd10	Homeo box D10	-1.91	0.007866	2.18
1417411_at	Nap1l5	Nucleosome assembly protein 1-like 5	-1.91	0.047023	-0.29
NuGO_emt034831_at	Nr2e3	Nuclear receptor subfamily 2, group E, member 3	-1.91	0.040798	-0.06
1434740_at	Scarf2	Scavenger receptor class F, member 2	-1.91	0.046195	-0.26
1420858_at	Pkia	Protein kinase inhibitor, alpha	-1.92	0.003616	3.21
1457072_at	Bcl11a	B-cell CLL/lymphoma 11A (zinc finger protein)	-1.93	0.04853	-0.35
1428347_at	Cyfp2	Cytoplasmic FMR1 interacting protein 2	-1.93	0.000484	5.87
1448823_at	Cxcl12	Chemokine (C-X-C motif) ligand 12	-1.95	0.007369	2.28
1436051_at	Myo5a	Myosin Va	-1.95	0.000596	5.61
1425065_at	Oas2	2'-5' oligoadenylate synthetase 2	-1.96	0.032484	0.34
1454876_at	Rab23	RAB23, member RAS oncogene family	-1.97	0.029208	0.48
NuGO_emt022150_at	Cartpt	CART prepropeptide	-1.98	0.008035	2.16
1423396_at	Agt	Angiotensinogen (serpin peptidase inhibitor, clade A, member 8)	-1.98	0.017774	1.15
1418213_at	Krt23	Keratin 23	-1.99	0.031284	0.38
1444670_at	Smyd3	SET and MYND domain containing 3	-2.02	0.022766	0.80
1453251_at	Dhx30	DEAH (Asp-Glu-Ala-His) box polypeptide 30	-2.03	0.027622	0.56
1440925_at	Rhoq	Ras homolog gene family, member Q	-2.05	0.031006	0.40
1422640_at	Podhb9	Protocadherin beta 9	-2.07	0.015304	1.34
1450708_at	Scg2	Secretogranin II	-2.09	0.010099	1.81
1435673_at	Hoxd3	Homeo box D3	-2.09	0.029426	0.47
1416710_at	Tmem35	Transmembrane protein 35	-2.10	0.007507	2.24
1423150_at	Scg5	Secretogranin V	-2.11	0.002438	3.69
1418392_a_at	Gbp3	Guanylate nucleotide binding protein 3	-2.11	0.020383	0.99
1436566_at	Rab40b	Rab40b, member RAS oncogene family	-2.12	0.003512	3.25
1441231_at	EG665123	Predicted gene, EG665123	-2.14	0.035789	0.15
1419349_a_at	Cyp2d9	Cytochrome P450, family 2, subfamily d, polypeptide 9	-2.14	0.003512	3.25
1445481_at	Al317158	Expressed sequence Al317158	-2.18	0.033108	0.26
1443698_at	Fbxo39	F-box protein 39	-2.19	0.009359	1.95
1424900_at	Slc29a4	Solute carrier family 29 (nucleoside transporters), member 4	-2.22	0.001105	4.77
1419185_a_at	Mlxip1	MLX interacting protein-like	-2.23	0.000867	5.09
1435504_at	Clip4	CAP-GLY domain containing linker protein family, member 4	-2.24	0.006244	2.51

1438868_at	Phf11	PHD finger protein 11	-2.24	0.000857	5.12
1422860_at	Nts	Neurotensin	-2.28	1.04E-09	19.02
1451860_a_at	Trim30	Tripartite motif protein 30	-2.28	0.038792	0.03
1434788_at	D930050A07Rik	RIKEN cDNA D930050A07 gene	-2.33	0.009359	1.95
1450684_at	Etv1	Ets variant gene 1	-2.38	0.000606	5.57
1433536_at	Lrp11	Low density lipoprotein receptor-related protein 11	-2.39	0.000793	5.21
NuGO_emt060551_at	9030421J09Rik	RIKEN cDNA 9030421J09 gene	-2.44	0.044497	-0.20
1428758_at	Tmem86a	Transmembrane protein 86A	-2.45	0.004961	2.76
1445881_at	NA	NA	-2.54	0.031006	0.39
1451426_at	D11Lgp2e	DNA segment, Chr 11, Lothar Hennighausen 2, expressed	-2.55	0.0093	1.98
1416639_at	Slc2a5	Solute carrier family 2 (facilitated glucose transporter), member 5	-2.67	0.006738	2.43
1429313_at	Ror1	Receptor tyrosine kinase-like orphan receptor 1	-2.70	0.001269	4.59
1433184_at	6720477C19Rik	RIKEN cDNA 6720477C19 gene	-2.72	0.017774	1.16
1419136_at	Akr1c18	Aldo-keto reductase family 1, member C18	-2.77	0.0113	1.66
1418113_at	Cyp2d10	Cytochrome P450, family 2, subfamily d, polypeptide 10	-2.79	0.000483	5.90
1417988_at	Resp18	Regulated endocrine-specific protein 18	-2.83	0.003512	3.26
1453196_a_at	Oasl2	2'-5' oligoadenylate synthetase-like 2	-2.84	0.011794	1.62
1423555_a_at	Ifi44	Interferon-induced protein 44	-2.86	0.042694	-0.13
1449939_s_at	Dlk1	Delta-like 1 homolog (Drosophila)	-2.99	5.20E-10	19.73
NuGO_emt091063_at	NA	NA	-3.08	0.000637	5.49
1436998_at	Ankrd43	Ankyrin repeat domain 43	-3.13	0.001056	4.85
1418293_at	Ifit2	Interferon-induced protein with tetratricopeptide repeats 2	-3.14	0.000637	5.44
NuGO_emt044299_at	Sstr1	Somatostatin receptor 1	-3.16	2.31E-08	16.00
1421492_at	Ptgds2	Prostaglandin D2 synthase 2, hematopoietic	-3.18	7.03E-05	8.13
1449025_at	Ifit3	Interferon-induced protein with tetratricopeptide repeats 3	-3.34	0.007674	2.21
1455528_at	NA	NA	-3.47	0.002046	3.94
1429273_at	Bmper	BMP-binding endothelial regulator	-3.55	0.000199	6.93
1425952_a_at	Gcg	Glucagon	-3.68	3.98E-09	17.68
1448628_at	Scg3	Secretogranin III	-3.77	1.61E-08	16.39
1448201_at	Sfrp2	Secreted frizzled-related protein 2	-4.18	0.000257	6.62
1418191_at	Usp18	Ubiquitin specific peptidase 18	-4.45	0.046848	-0.28
1449493_at	InsI5	Insulin-like 5	-9.54	5.98E-16	30.51
1419473_a_at	Cck	Cholecystokinin	-11.97	1.17E-11	23.32

Table S4C. Transcripts differentially expressed in the ileum at 14d between *R. hominis* inoculated animals and germfree animals.

ID	Symbol	Name	FC	P-value	B
1427343_at	Rasd2	RASD family, member 2	4.21	0.008352	3.02
1420673_a_at	Acox2	Acyl-Coenzyme A oxidase 2, branched chain	3.37	0.023445	1.29
1418174_at	Dbp	D site albumin promoter binding protein	3.26	0.04708	0.16
1434116_at	Cbx2	Chromobox homolog 2 (Drosophila Pc class)	3.15	0.01298	2.21
1456284_at	Tmem171	Transmembrane protein 171	2.94	0.011378	2.45
1460713_at	BC048355	cDNA sequence BC048355	2.61	0.035037	0.67
1438689_at	4632433K11Rik	RIKEN cDNA 4632433K11 gene	2.41	0.04048	0.41
1460187_at	Sfrp1	Secreted frizzled-related sequence protein 1	2.37	0.021033	1.51
1420645_at	Pcgef2	Polycomb group ring finger 2	2.31	0.003987	4.19
1455547_at	Zc3h7b	Zinc finger CCH type containing 7B	2.16	0.008352	2.98
1416258_at	Tk1	Thymidine kinase 1	2.12	0.017451	1.78
1449845_a_at	Ephb4	Eph receptor B4	2.12	0.033002	0.77
1455120_at	Hpdl	4-hydroxyphenylpyruvate dioxygenase-like	2.07	0.023445	1.30
1417399_at	Gas6	Growth arrest specific 6	2.06	0.000167	8.81
1452862_at	Rreb1	Ras responsive element binding protein 1	2.05	0.03901	0.47
1455246_at	NA	NA	2.02	0.032868	0.78
1434322_at	Micall2	MICAL-like 2	1.99	0.009875	2.69
1428207_at	Bcl7a	B-cell CLL/lymphoma 7A	1.99	0.04708	0.11
1420845_at	Mrps2	Mitochondrial ribosomal protein S2	1.97	0.042267	0.36
1444254_at	NA	NA	1.96	0.025132	1.19
1448656_at	Cacnb3	Calcium channel, voltage-dependent, beta 3 subunit	1.96	0.032019	0.86
1434908_at	AI480556	Expressed sequence AI480556	1.95	0.029591	0.99
1424376_at	Cdc42ep1	CDC42 effector protein (Rho GTPase binding) 1	1.89	0.01298	2.20
1430274_a_at	Stard3nl	STARD3 N-terminal like	1.88	0.008596	2.92
1416513_at	Lamb2	Laminin, beta 2	1.87	0.016398	1.89
1416536_at	Mum1	Melanoma associated antigen (mutated) 1	1.86	0.025132	1.17
NuGO_empt084041_s_at	Defcr20	Defensin related cryptdin 20	1.84	0.020502	1.55
1418320_at	Prss8	Protease, serine, 8 (proctasin)	1.83	0.021033	1.49
1455163_at	Guf1	GUF1 GTPase homolog (S. cerevisiae)	1.80	0.02461	1.22
1436665_a_at	Ltbp4	Latent transforming growth factor beta binding protein 4	1.80	0.036581	0.57
1420643_at	Lfng	Lunatic fringe gene homolog (Drosophila)	1.79	0.003248	4.57

1428695_at	9130227C08Rik	RIKEN cDNA 9130227C08Rik gene	1.79	0.036167	0.59
1453018_at	Nvl	Nuclear VCP-like	1.77	0.035037	0.65
1419101_at	Sin3a	Transcriptional regulator, SIN3A (yeast)	1.76	0.030882	0.91
1424459_at	Aylt2	Acyltransferase like 2	1.74	0.039921	0.43
1415935_at	Smoc2	SPARC related modular calcium binding 2	1.73	0.04615	0.19
1424618_at	Hpd	4-hydroxyphenylpyruvic acid dioxygenase	1.73	0.03901	0.47
1426297_at	Tcf2a	Transcription factor E2a	1.73	0.002215	5.11
1425391_a_at	Osbp15	Oxysterol binding protein-like 5	1.73	0.04615	0.19
1455719_at	Tubb5	Tubulin, beta 5	1.73	0.032477	0.82
1451912_a_at	Fgfr1	Fibroblast growth factor receptor-like 1	1.73	0.02321	1.33
1429582_at	Btd14a	BTB (POZ) domain containing 14A	1.72	0.003987	4.21
1454777_at	Slco2b1	Solute carrier organic anion transporter family, member 2b1	1.70	0.01298	2.22
1424101_at	Hnrpl	Heterogeneous nuclear ribonucleoprotein L	1.69	0.044481	0.25
1448691_at	Ubqln4	Ubiquilin 4	1.65	0.016054	1.92
1417604_at	Camk1	Calcium/calmodulin-dependent protein kinase I	1.63	0.033233	0.74
1442757_at	Lrch1	Leucine-rich repeats and calponin homology (CH) domain containing 1	1.63	0.007505	3.30
1460675_at	Igsf8	Immunoglobulin superfamily, member 8	1.62	0.04708	0.13
1418671_at	Capn5	Calpain 5	1.58	0.04708	0.15
1426897_at	Rcc2	Regulator of chromosome condensation 2	1.57	0.035037	0.65
1417594_at	Gkap1	G kinase anchoring protein 1	1.52	0.019831	1.60
1433539_at	Comm3	COMM domain containing 3	1.51	0.009875	2.72
1435469_at	Qscn6l1	Quiescin Q6-like 1	1.51	0.041412	0.38
1448561_at	Ncf2	Neutrophil cytosolic factor 2	1.51	0.044087	0.30
1427022_at	Ddx42	DEAD (Asp-Glu-Ala-Asp) box polypeptide 42	1.50	0.01872	1.67
1416030_a_at	Mcm7	Minichromosome maintenance deficient 7 (S. cerevisiae)	1.49	0.02321	1.34
1450023_at	Gtbbp1	GTP binding protein 1	1.48	0.044113	0.27
1417879_at	Nenf	Neuron derived neurotrophic factor	1.48	0.03901	0.46
1424640_at	Arl8a	ADP-ribosylation factor-like 8A	1.47	0.044113	0.28
1418982_at	Cebpa	CCAAT/enhancer binding protein (C/EBP), alpha	1.47	0.036167	0.60
1428382_at	Smrcc2	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 2	1.47	0.044113	0.28
1434134_at	Wdr42a	WD repeat domain 42A	1.45	0.009875	2.74
1450519_a_at	Prkca	Protein kinase, cAMP dependent, catalytic, alpha	1.44	0.008352	2.99

1451306_at	Cdca7l	Cell division cycle associated 7 like	1.44	0.035037	0.65
1426724_at	Cnn3	Calponin 3, acidic	1.44	0.033105	0.76
1424644_at	Tbcc	Tubulin-specific chaperone c	1.42	0.032128	0.85
1417266_at	Ccl6	Chemokine (C-C motif) ligand 6	1.42	0.021206	1.47
1415975_at	Carhsp1	Calcium regulated heat stable protein 1	1.38	0.014132	2.10
1448277_at	Pold2	Polymerase (DNA directed), delta 2, regulatory subunit	1.38	0.044481	0.24
1433736_at	Hcfc1	Host cell factor C1	1.35	0.035162	0.64
1435149_at	Plcg1	Phospholipase C, gamma 1	1.35	0.036167	0.59
1417500_a_at	Tgm2	Transglutaminase 2, C polypeptide	1.33	0.022457	1.39
1428125_at	4921506J03Rik	RIKEN cDNA 4921506J03 gene	1.32	0.030752	0.92
1452100_at	Dullard	Dullard homolog (Xenopus laevis)	1.32	0.030391	0.94
1448148_at	Grn	Granulin	1.30	0.011783	2.33
1451984_at	Hnrp1l	Heterogeneous nuclear ribonucleoprotein U-like 1	1.30	0.016884	1.85
1426401_at	Ppp3ca	Protein phosphatase 3, catalytic subunit, alpha isoform	1.29	0.009875	2.72
1428380_at	0610007C21Rik	RIKEN cDNA 0610007C21 gene	1.26	0.043354	0.32
1418364_a_at	Ftl1	Ferritin light chain 1	1.24	0.01702	1.82
1456854_at	Neurl	Neuralized-like homolog (Drosophila)	-1.18	0.031207	0.89
1448853_at	Synj2bp	Synaptojanin 2 binding protein	-1.19	0.04708	0.15
1416281_at	Wdr45l	Wdr45 like	-1.22	0.016971	1.84
1418843_at	Slc30a4	Solute carrier family 30 (zinc transporter), member 4	-1.30	0.037239	0.54
1459557_at	Zbtb16	Zinc finger and BTB domain containing 16	-1.31	0.001423	5.62
1418116_at	Ifrg15	Interferon alpha responsive gene	-1.32	0.038411	0.50
1448762_at	Rad17	RAD17 homolog (S. pombe)	-1.32	0.044087	0.30
1435461_at	Magi3	Membrane associated guanylate kinase, WW and PDZ domain containing 3	-1.33	0.044113	0.26
1434835_at	Wapal	Wings apart-like homolog (Drosophila)	-1.33	0.046851	0.17
1444328_at	NA	NA	-1.34	0.010408	2.57
NuGO_empt073151_at	Nlrp9b	NLR family, pyrin domain containing 9B	-1.34	0.045959	0.20
1427269_at	Sfrs11	Splicing factor, arginine/serine-rich 11	-1.35	0.02461	1.22
1436157_at	Ccar1	Cell division cycle and apoptosis regulator 1	-1.36	0.021206	1.46
NuGO_empt081039_at	Eif4e1b	Eukaryotic translation initiation factor 4E family member 1B	-1.38	0.037239	0.54
1422217_a_at	Cyp1a1	Cytochrome P450, family 1, subfamily a, polypeptide 1	-1.38	0.005253	3.71
1434654_at	Cog3	Component of oligomeric golgi complex 3	-1.38	0.03901	0.46
1421680_at	NA	NA	-1.39	0.044113	0.28

1424296_at	Gclc	Glutamate-cysteine ligase, catalytic subunit	-1.41	0.023137	1.36
1440722_at	D19Erttd386e	DNA segment, Chr 19, ERATO Doi 386, expressed	-1.41	0.008352	3.15
1429849_at	4632411B12Rik	RIKEN cDNA 4632411B12 gene	-1.44	0.015927	1.94
1451407_at	Jam4	Junction adhesion molecule 4	-1.44	0.015521	1.99
1424324_at	Esco1	Establishment of cohesion 1 homolog 1 (S. cerevisiae)	-1.47	0.010277	2.60
1441403_at	6430501K19Rik	RIKEN cDNA 6430501K19 gene	-1.47	0.019952	1.59
1453160_at	Thrap1	Thyroid hormone receptor associated protein 1	-1.48	0.034193	0.71
1432962_at	2610024D14Rik	RIKEN cDNA 2610024D14 gene	-1.49	0.018436	1.70
1456896_at	6720462K09Rik	RIKEN cDNA 6720462K09 gene	-1.50	0.011783	2.34
1444705_at	App	Amyloid beta (A4) precursor protein	-1.50	0.032477	0.80
1426886_at	Cln5	Ceroid-lipofuscinosis, neuronal 5	-1.52	0.01702	1.81
1459059_at	2010308F09Rik	RIKEN cDNA 2010308F09 gene	-1.53	0.008352	2.97
1436616_at	R74740	Expressed sequence R74740	-1.55	0.011378	2.44
1453269_at	Unc5b	Unc-5 homolog B (C. elegans)	-1.55	0.018436	1.70
1424536_at	Oas1e	2'-5' oligoadenylate synthetase 1E	-1.55	0.008352	3.08
1444565_at	NA	NA	-1.56	0.005253	3.70
1448049_at	Jmjd1c	Jumonji domain containing 1C	-1.58	0.02321	1.33
1441546_at	Mpp6	Membrane protein, palmitoylated 6 (MAGUK p55 subfamily member 6)	-1.58	0.009875	2.70
1442605_at	Bach2	BTB and CNC homology 2	-1.59	0.032477	0.80
1451415_at	1810011O10Rik	RIKEN cDNA 1810011O10 gene	-1.66	0.023445	1.29
1436637_at	Elf4h	Eukaryotic translation initiation factor 4H	-1.66	0.011571	2.40
1453457_at	Sri	Sorcin	-1.67	0.024152	1.25
1429680_at	Tra2a	Transformer 2 alpha homolog (Drosophila)	-1.68	0.033105	0.75
1429624_at	Sltm	SAFB-like, transcription modulator	-1.68	0.036906	0.56
1429870_at	Tnik	TRAF2 and NCK interacting kinase	-1.72	0.001356	5.94
1444065_at	Cyb5d2	Cytochrome b5 domain containing 2	-1.72	0.000237	8.03
1424208_at	Ptger4	Prostaglandin E receptor 4 (subtype EP4)	-1.74	0.029592	0.98
1452837_at	Lpin2	Lipin 2	-1.74	0.011783	2.33
1448185_at	Herpud1	Homocysteine-inducible, endoplasmic reticulum stress-inducible, ubiquitin-like domain member 1	-1.75	0.002895	4.81
1432719_at	4833412K13Rik	RIKEN cDNA 4833412K13 gene	-1.76	0.04708	0.13
1437868_at	BC023892	cDNA sequence BC023892	-1.76	0.044113	0.28
1433101_at	9030419F21Rik	RIKEN cDNA 9030419F21 gene	-1.77	0.008596	2.90

1451621_at	5830417C01Rik	RIKEN cDNA 5830417C01 gene	-1.77	0.00495	3.81
1432423_a_at	C530008M17Rik	RIKEN cDNA C530008M17 gene	-1.78	0.01164	2.38
1429399_at	Rnf125	Ring finger protein 125	-1.84	0.033002	0.77
1453264_at	Marveld3	MARVEL (membrane-associating) domain containing 3	-1.86	4.82E-05	10.38
1454343_at	Ppapdc1	Phosphatidic acid phosphatase type 2 domain containing 1	-1.88	0.003658	4.42
1435571_at	A530065I17Rik	RIKEN cDNA A530065I17 gene	-1.90	0.015795	1.96
1443164_at	NA	NA	-1.91	0.010163	2.63
1437776_at	Tmcc1	Transmembrane and coiled coil domains 1	-1.91	0.028192	1.04
1442111_at	D430033H22Rik	RIKEN cDNA D430033H22 gene	-1.91	0.007072	3.40
1424451_at	Acaa1b	Acetyl-Coenzyme A acyltransferase 1B	-1.92	0.035037	0.67
1459879_at	4921513D23Rik	RIKEN cDNA 4921513D23 gene	-1.92	0.008352	2.99
1428776_at	Slc10a6	Solute carrier family 10 (sodium/bile acid cotransporter family), member 6	-1.92	0.011378	2.45
1458079_at	Usp40	Ubiquitin specific peptidase 40	-1.93	0.004696	3.98
1442897_at	2610024E20Rik	RIKEN cDNA 2610024E20 gene	-1.95	0.00402	4.15
1455744_at	Tmem181	Transmembrane protein 181	-1.97	0.037843	0.52
1449385_at	Hsd17b6	Hydroxysteroid (17-beta) dehydrogenase 6	-1.98	0.025132	1.18
1443068_at	D130084N16Rik	RIKEN cDNA D130084N16 gene	-1.98	0.04708	0.11
1419582_at	Cyp2c55	Cytochrome P450, family 2, subfamily c, polypeptide 55	-1.99	0.042267	0.35
1457801_at	9930024M15Rik	RIKEN cDNA 9930024M15 gene	-2.01	0.04708	0.13
1438331_at	Ypel2	Yippee-like 2 (Drosophila)	-2.02	0.037843	0.51
1428833_at	4930406D14Rik	RIKEN cDNA 4930406D14 gene	-2.04	0.015418	2.01
1419388_at	Tm4sf20	Transmembrane 4 L six family member 20	-2.05	0.010163	2.65
1455510_at	Spop	Speckle-type POZ protein	-2.08	0.023445	1.29
1444178_at	ENSMUSG000000052976	Predicted gene, ENSMUSG000000052976	-2.09	0.029592	0.97
1443159_at	Txnrd1	Thioredoxin reductase 1	-2.10	0.04708	0.11
1457161_at	9530029O12Rik	RIKEN cDNA 9530029O12 gene	-2.11	0.019032	1.65
1459887_at	NA	NA	-2.11	0.003987	4.29
1459005_at	NA	NA	-2.13	0.007384	3.34
1439283_at	NA	NA	-2.18	0.007823	3.24
1438338_at	Mdh1	Malate dehydrogenase 1, NAD (soluble)	-2.21	0.001446	5.54
1458452_at	NA	NA	-2.21	0.010277	2.59
1442256_at	Prkcd	Protein kinase C, delta	-2.21	0.026633	1.11
1443969_at	Irs2	Insulin receptor substrate 2	-2.22	0.003987	4.22

1452462_a_at	Banp	Btg3 associated nuclear protein	-2.30	0.021033	1.52
1454558_at	5430416B10Rik	RIKEN cDNA 5430416B10 gene	-2.32	1.71E-07	15.68
1430393_at	C030048B08Rik	RIKEN cDNA C030048B08 gene	-2.32	0.001388	5.78
1446929_at	Bach2	BTB and CNC homology 2	-2.32	0.027369	1.07
NuGO_empt067737_at	9130230L23Rik	RIKEN cDNA 9130230L23 gene	-2.34	0.04708	0.15
1441138_at	Foxn2	Forkhead box N2	-2.36	0.001356	6.01
1416041_at	Sgk	Serum/glucocorticoid regulated kinase	-2.40	0.04048	0.41
1454158_at	Mpp7	Membrane protein, palmitoylated 7 (MAGUK p55 subfamily member 7)	-2.40	0.011571	2.40
1459253_at	Arddc3	Arrestin domain containing 3	-2.46	0.044113	0.27
1444376_at	Sesn1	Sestrin 1	-2.48	0.036167	0.59
1430362_at	5730409N24Rik	RIKEN cDNA 5730409N24 gene	-2.49	0.032477	0.83
1442069_at	D5Wsu178e	DNA segment, Chr 5, Wayne State University 178, expressed	-2.50	0.032477	0.80
1433203_at	6030400A10Rik	RIKEN cDNA 6030400A10 gene	-2.56	0.000747	6.76
1456706_at	1700109H08Rik	RIKEN cDNA 1700109H08 gene	-2.60	0.004825	3.89
1441561_at	Fbxl3	F-box and leucine-rich repeat protein 3	-2.60	0.034193	0.70
1437884_at	Arl5b	ADP-ribosylation factor-like 5B	-2.65	0.022457	1.40
1445843_at	Chd2	Chromodomain helicase DNA binding protein 2	-2.66	0.044481	0.24
1428306_at	Ddit4	DNA-damage-inducible transcript 4	-2.66	0.008352	3.09
1421009_at	Rsad2	Radical S-adenosyl methionine domain containing 2	-2.69	0.04708	0.14
1440227_at	BF642829	Expressed sequence BF642829	-2.69	0.026633	1.11
1453595_at	Bcl6	B-cell leukemia/lymphoma 6	-2.70	0.004825	3.89
1444775_at	9930033D15Rik	RIKEN cDNA 9930033D15 gene	-2.71	0.003248	4.61
1439972_at	Etnk1	Ethanolamine kinase 1	-2.74	0.032477	0.83
1440536_at	Slc22a5	Solute carrier family 22 (organic cation transporter), member 5	-2.90	0.021033	1.49
1438660_at	Gcnt2	Glucosaminyl (N-acetyl) transferase 2, l-branching enzyme	-2.95	0.018436	1.70
1437759_at	Pfkfb	Phosphofructokinase, platelet	-2.97	0.008352	3.11
NuGO_empt050020_at	Amica1	Adhesion molecule, interacts with CXADR antigen 1	-2.97	0.000981	6.41
1446950_at	Tox	Thymocyte selection-associated HMG box gene	-3.00	0.008352	2.98
1447141_at	AW107722	Expressed sequence AW107722	-3.14	0.013251	2.17
1459219_at	Rapgef2	Rap guanine nucleotide exchange factor (GEF) 2	-3.21	0.000747	6.75
1449496_at	2010109I03Rik	RIKEN cDNA 2010109I03 gene	-3.22	0.000167	8.69
1458296_at	NA	NA	-3.23	0.001388	5.70
1440749_at	NA	NA	-3.38	0.001388	5.76

1440892_at	BC017647	cDNA sequence BC017647	-3.52	0.004881	3.85
1441115_at	Rnf125	Ring finger protein 125	-3.90	0.000237	8.14
1421365_at	Fst	Follistatin	-4.33	0.04708	0.12
1446972_at	D15Wsu126e	DNA segment, Chr 15, Wayne State University 126, expressed	-7.09	0.003054	4.72

Table S4D. Transcripts differentially expressed in the ileum at 28d between *R. hominis* inoculated animals and germfree animals.

ID	Symbol	Name	M	P-value	B
1458427_at	Brip1	BRCA1 interacting protein C-terminal helicase 1	-1.72	0.002194	4.97

5 Genes involved in innate immunity and gut barrier function were significantly induced by the presence of *R. hominis* in the ascending colon. The GO-process 'innate immune response' (GO:0045087) was up-regulated and included the important TLR-related genes *Tlr5*, *Tlr1* and *Vnn1*. The up-regulation of *Tlr5* was of particular interest, given the corresponding induction of flagellar genes and flagellin protein in *R. hominis* during gut colonization, and may infer a role for this innate signalling pathway in mediating innate and adaptive immune responses.

10 Other innate immune genes affected in the colon by *R. hominis* included the antimicrobial peptides *Defb37*, *Pla2g3*, *Muc16* and *Itln* and the gut barrier function genes *Spr1a*, *Cldn4*, *Pmp22*, *Crb3* and *Magi3*. Innate immune genes showing up-regulation in the ileum in response to *R. hominis* included *Defcr20*, *Pcgf2*, *Ltbp4*, *Igsf8* and *Tcfe2a*. Interestingly,

15 herein it is shown negative regulation of the NF- κ B pathway (GO:0043124) (Fig. S2) by *R. hominis*, which, like *B. thetaiotaomicron* (Kelly *et al.* 2004), may contribute to immune homeostasis by down-regulating this inflammatory cascade.

To demonstrate that these responses are specific to *R. hominis*, the response of germfree animals to another commensal bacterium was investigated. The gene expression responses to colonization with *E. coli* MG1655 (K12) were compared to *R. hominis* after 10-14d and 22-28d post-colonization. Over this time interval, large differences in gene expression were observed in response to *R. hominis* but not to *E. coli*, indicating that *R. hominis* is biologically very active in the gut, in contrast to the minimal impact of *E. coli* (Fig. S3). The response to

20 *E. coli* deduced from the gene expression data in the ascending colon seemed to be mainly a B-cell mediated antibody response.

***R. hominis* affects T cell pathways mostly in the colon**

30 The majority of pathways affected at 14d grouped into the categories cell differentiation, cell cycle regulation and tissue remodelling. Importantly, immune response was a major pathway induced at 28d in the ascending colon. The significantly affected pathways in this category were mostly involved in T cell function, including IL-10 signalling and regulation of T cell function by CTLA-4 (Table S5). The genes involved in these pathways showed both up-

35 regulation and down-regulation, so while these pathways were significantly affected by the

presence of *R. hominis*, the precise net functional effects on T cell differentiation required further investigation. To clarify the role of *R. hominis* in relation to T cell differentiation, conventional mice with *R. hominis* were treated for 14 days and determined the impact on T cell subsets both in the lamina propria and mesenteric lymph nodes (MLN). Treatment with *R. hominis* increased the population of CD3⁺CD4⁺CD25⁺FoxP3⁺ T cells in both locations (Fig. 4). An assessment was made of the numbers of double-positive CD3⁺FoxP3⁺ cells in the lamina propria of ascending and descending colon of mono-associated C3H/HeN and C57Bl6 animals and confirmed a significant increase in regulatory T cells in *R. hominis*-treated mice (Fig. 5A). The GO-process for 'actin polymerization' (GO:0030041) (*Arpc3*, *Capg*, *Cdc42ep5* and *Rhoc*) was up-regulated at 28d in the colon in *R. hominis* colonized mice (Fig. S2). Actin polymerization at the immune synapse is required for T cell activation and effector function. Overall, this data indicates that *R. hominis* actively effects the adaptive immune response in the colon by positively influencing T cell regulation.

Table S5. Immune system response pathway analysis of transcripts differentially expressed in the ascending colon between *R. hominis* treated mice and germfree mice at 28d. Differentially expressed genes ($P < 0.05$) were imported into GeneGo MetaCore analytical software to determine significantly enriched canonical pathways in each group. * The number of genes on each map that are differentially expressed in the specific treatment comparison. ** The total number of genes on each map.

Regulatory processes/Immune system response pathway	P-value	Significant*	Total**
Immune response_IL-10 signaling pathway	0.00125	10	26
Immune response_IL-9 signaling pathway	0.00592	11	36
Immune response_HMGB1/RAGE signaling pathway	0.00832	14	53
Immune response_BCR pathway	0.00992	14	54
Development_GM-CSF signaling	0.01258	13	50
Development_PEDF signaling	0.02618	12	49
Immune response_IL-5 signalling	0.02840	11	44
Immune response_TCR and CD28 co-stimulation in activation of NF-kB	0.03611	10	40
Immune response_Regulation of T cell function by CTLA-4	0.04598	9	36
Immune response_CD40 signaling	0.04796	14	65
Signal transduction_JNK pathway	0.04921	10	42

Related to these results was the induction of the Ly6 gene family in the ascending colon. In particular, the GPI-anchored gene product of *Ly6g6c* was up-regulated 25-fold, and the related gene *Ly6g6e* was up-regulated two-fold at 28d. Most haematopoietic cells, including neutrophils and plasmacytoid dendritic cells, express one or more members of the Ly6 family. Furthermore, a possible role of Ly6 in T cell activation, differentiation and maturation has been proposed (Mallya, Campbell & Aguado 2006). Immunocytochemistry confirmed increased presence of Ly6G⁺, CD11b⁺ and CD3⁺ FoxP3⁺ cells in *R. hominis*-colonized mice (Fig. 5B).

***R. hominis* flagellins modulate T cell differentiation**

The influence of bacteria on the differentiation of T cells may reflect the array of TLR ligands displayed. For example, the coupling between TLR5 signalling and CD4⁺ T cell responses has recently been demonstrated for flagellate pathogens (Letran *et al.* 2011). Interestingly, depending upon the experimental setting, flagellin can prime a range of T cell responses including Th1, Th2, Th17 and Treg responses (Wilson *et al.* 2012).

The functionality of bacterial flagellins FlaA1 (RH1) and FlaA2 (RH2) was investigated using novel soluble recombinant flagellin proteins, generated against the unique *R. hominis* flagellin sequences. The ability of RH1 and RH2 were compared and contrasted with various commensal and pathogenic flagellins, generated using identical protocols, to activate signalling responses in intestinal epithelial cell lines and bone marrow derived dendritic cells

expanded with either FLT3L or GM-CSF.

Epithelial cells treated with identical concentrations of different bacterial flagellins revealed distinct patterns of gene expression (Fig. 6A). Importantly, no endotoxin contamination was detected in the recombinant protein preparations. *Salmonella enteritidis* (SE) was more potent than *E. coli* K12 (EC) or RH1 flagellin. RH1 flagellin also showed a strong response but clustered in a distinct clad along with commensal EC. The responses were shown to be TLR5-dependent using epithelial cells expressing dominant-negative TLR5. In contrast, RH2 was shown to be minimally active; it was generally not pro-inflammatory nor did it activate the conserved gene signature (IL-8, CXCL-1, CXCL-2 and CXCL-10) induced by other recombinant bacterial flagellins. RH1 flagellin protein is more biologically active than RH2 *in vitro*; although both recombinant proteins were expressed *in vivo*, RH1 was also significantly up-regulated at the gene expression level *in vivo*. It was demonstrated that the RH1 flagellin from *R. hominis* induced different responses in Flt3L and GM-CSF derived dendritic cells relative to commensal *E. coli* and pathogenic *Salmonella enteritidis* (Fig. 6B-C). In particular, RH1 was unique in its ability to activate Flt3L-expanded DCs, with up-regulation of I-A/I-E and CD40 and the production of IL-10 by bone marrow derived DCs from both C3H/HeN and C57Bl/6 mice. The IL-10/IL-12 ratio was particularly elevated in C57Bl/6 DCs (Fig. 6D), which were found to be CD103⁺ Siglec-H⁺. Consistent with the observations herein, a number of recent reports have also shown that flagellin can activate CD103⁺DC populations (Flores-Langarica *et al.* 2012, Kinnebrew *et al.* 2012).

To evaluate the functional importance of *R. hominis* and its flagellins, germ-free TLR5KO and WT mice were mono-colonized. The heatmap showing differentially expressed genes for both TLR5KO and wild-type colonized with *R. hominis* revealed a very strong effect of TLR5 (Fig S4). Although T cell pathways were still influenced by *R. hominis* colonization in TLR5KO mice, the responses were more related to IL4, IL5, IL-6, IL-9 pathways and not IL-10 and CTLA-4 (Table S6). Furthermore, the numbers of double-positive CD3⁺FoxP3⁺ cells in the lamina propria of TLR5KO mice were not increased by *R. hominis* treatment (Fig. 6E), in contrast to mono-associated C3H/HeN and C57Bl6 animals (Fig. 5A).

Table S6. Immune system response pathway analysis of transcripts differentially expressed in the ascending colon between TLR5 KO mice and WT mice, mono-colonized with *R. hominis*. Differentially expressed genes ($P<0.05$) were imported into GeneGo MetaCore analytical software to determine significantly enriched canonical pathways in each group. * The number of genes on each map that are differentially expressed in the specific treatment comparison. ** The total number of genes on each map.

#	Maps	pValue	Ratio
1	Immune response_IL-9 signaling pathway	0.00001	7 36
2	Immune response_Histamine signaling in dendritic cells	0.00008	7 50
3	Immune response_HMGB1/RAGE signaling pathway	0.00012	7 53
4	Immune response_IL-6 signaling pathway	0.00046	5 31
5	Immune response_Histamine H1 receptor signaling in immune response	0.00052	6 48
6	Immune response_Oncostatin M signaling via MAPK in mouse cells	0.00082	5 35
7	Immune response_Oncostatin M signaling via MAPK in human cells	0.00107	5 37
8	Signal transduction_JNK pathway	0.00191	5 42
9	Immune response_IL-7 signaling in B lymphocytes	0.00213	5 43
10	Immune response_Signaling pathway mediated by IL-6 and IL-1	0.00362	4 30
11	Development_GM-CSF signaling	0.00415	5 50
12	Immune response_T cell receptor signaling pathway	0.00492	5 52
13	Chemotaxis_Leukocyte chemotaxis	0.00528	6 75
14	Immune response_CCL2 signaling	0.00579	5 54
15	Immune response_CD28 signaling	0.00579	5 54
16	Immune response_Role of DAP12 receptors in NK cells	0.00579	5 54
17	Immune response_Fc epsilon RI pathway	0.00626	5 55
18	Immune response_Role of PKR in stress-induced antiviral cell response	0.00728	5 57
19	Immune response_HMGB1 release from the cell	0.01114	4 41
20	Immune response_IL-15 signaling	0.01176	5 64
21	Immune response_HTR2A-induced activation of cPLA2	0.01313	4 43
22	Immune response_IL-4 signaling pathway	0.01421	4 44
23	Immune response_IL-5 signalling	0.01421	4 44
24	Immune response_Fc gamma R-mediated phagocytosis in macrophages	0.01653	4 46
25	Immune response_NF-AT signaling and leukocyte interactions	0.01653	4 46
26	Development_PEDF signaling	0.02043	4 49
27	Immune response_IL-2 activation and signaling pathway	0.02043	4 49
28	Immune response_NFAT in immune response	0.02332	4 51
29	Immune response_IL-3 activation and signaling pathway	0.02803	3 31
30	Immune response_BCR pathway	0.02809	4 54
31	Immune response_TLR signaling pathways	0.02809	4 54
32	Immune response_Immunological synapse formation	0.03727	4 59
33	Immune response_Th17 cell differentiation	0.03836	3 35
34	Immune response_Human NKG2D signaling	0.04722	3 38

The observation herein that *R. hominis* influences Tregs in conventional and germ-free mice and not in TLR5KO is consistent with reports of flagellin–TLR5 interactions promoting CD4⁺CD25⁺Foxp3⁺ regulatory T cells (Crellin *et al.* 2005, Hossain *et al.* 2011). Similarly, the ability of a flagellin-ovalbumin fusion protein to suppress IL-4 production by OVA-T-cell receptor CD4(+) T cells via an IL-10 dependent mechanism has recently been described (Schulke *et al.* 2011) indicating that flagellin can influence the directional differentiation of T cell subsets. In addition, the impact of the TLR5KO on T cells responses driven by *R. hominis* infers that RH1 (the signalling flagellin) was crucial in mediating the Treg responses and not RH2 (the non-signaling flagellin). Finally, an additional observation was the

enhancement of Type I IFN genes in TLR5KO mice (including *Ifi202b*, *Ifi203* and *Irf4*), which suggests that TLR5 signalling may dampen Type I interferon responses.

***R. hominis* modulates innate immune response genes in both the ileum and colon and attenuates colitis in DSS-treated mice**

The DSS mouse model was used to test the therapeutic efficacy of *R. hominis*, due to the control of inflammatory pathways as well as the positive effects on Treg induction in mono-associated mice. Mice were dosed (~50 µL, 10⁹CFU) daily for a period of 14 days, and given DSS (MW 50kDa, 30g/l) in their drinking water from day 8 onwards. Gene expression of a panel of pro-inflammatory biomarkers showed that untreated DSS mice had strong elevation of all investigated genes compared to wild-type mice, with gene induction ranging from 4- to 49-fold (Fig. 7A). Pro-inflammatory gene induction was significantly lower in *R. hominis*-treated compared to untreated mice, indicating strong therapeutic benefits of oral administration of *R. hominis*. Histological analysis showed the presence of severe inflammation in the ascending colon of untreated DSS, while the colonic mucosa of *R. hominis*-treated animals was normal, with low-level inflammation, consistent with the reduced inflammatory gene expression (Fig. 7B and C).

***R. hominis* colonization influences body composition and expression of satiety genes**

Significant metabolic actions of *R. hominis* in mono-associated mice were also evident. The GO-processes 'negative regulation of response to food' (GO:0032096), 'negative regulation of appetite' (GO:0032099), and 'regulation of catecholamine secretion' (GO:0050433) were all down-regulated in the ascending colon after colonization with *R. hominis* (Fig. S5). This data infers that *R. hominis* exerts a stimulatory effect on host appetite. The genes involved in these processes were *Agt*, *Cartpt*, *Cck* and *Cxcl12*, with fold-changes ranging from 2- to 12-fold. *Cck*, in particular, plays a major role in digestion and satiety as a hunger suppressant. *Gcg* also showed down-regulation at this gut site.

To establish whether these gene changes had physiological relevance in relation to food intake and body composition, dry carcass weight and composition analyses were performed. The dry carcass weights of *R. hominis*-associated mice were significantly heavier compared to GF animals, and the differences were most discernible at 14d (Fig. S6A). Further carcass lipid analysis showed that total adiposity was also significantly higher in *R. hominis*-treated animals at 14d (Fig. S6B). These findings are consistent with recent data revealing the role of Firmicutes in energy harvest through dietary fermentation, but also support the notion that gut bacteria can in fact modulate the brain-gut axis and appetite-regulating hormones.

Discussion

The long-term co-evolution of host-microbe mutualism has likely driven the selection of functionally important bacterial species in the gut, the majority of which are not highly represented in other ecosystems. Currently, there is limited information regarding the contribution of individual members of the microbial community to intestinal functions, particularly in relation to development of the mucosal immune system.

Recent work using a reversible colonization model based on *E. coli* (HA 107) has demonstrated that live bacteria are required in numbers approaching 10^8 CFUs per gram of content for the immune-inducing effects on IgA (Hapfelmeier *et al.* 2010). The specific functions of SFB and *Bacteroides fragilis* have been investigated in the mouse gut to define their individual contributions to T cell biology and both these bacteria have been shown to be potent inducers of Tregs and Th17 cells (Mazmanian *et al.* 2005, Gaboriau-Routhiau *et al.* 2009, Ivanov *et al.* 2009). The effects of individual members of the cluster XIVa Firmicutes have not been reported previously, although their presence in the ASF and the contribution of a mixed culture of 46 Clostridial strains, which also affects T cell differentiation has been noted (Geuking *et al.* 2011, Atarashi *et al.* 2011).

Reported here is the first successful mono-association of the germ-free mouse gut with an anaerobic bacterium, *R. hominis*, which is a member of the Firmicutes phylum. The extreme oxygen sensitivity of bacteria like *Roseburia* requires strict anaerobic culture techniques, making it difficult to perform functional characterization. The stable mono-colonization of *R. hominis* in germfree mice has been established and the complete annotated genomic sequence to uncover its metabolic organization, physiology, and symbiotic properties has been produced. It was found that the transcriptional responses of *R. hominis* following colonization could be attributed to both the host gut environment and diet. The host-driven effects dominated the response of *R. hominis* following mono-association. These included gene transfer, membrane transport, chemotaxis and motility subsystems. The strong up-regulation of genes involved in mobilization transfer supports the view that the gut environment is highly conducive to horizontal gene exchange between members of the gut microbiota. Thus, this environment may accelerate the dissemination of genes important for bacterial survival, colonization and function within the gut ecosystem. The role of motility and flagellar apparatus in host colonization is well-elaborated for pathogenic bacteria but much less is known about the role of flagellar proteins in commensal bacteria. *In vivo* experiments revealed a stimulatory effect of the host intestinal environment on the expression of flagellin

genes.

A clear role for *R. hominis* in promoting gut barrier function and innate immunity in the mouse colon has been established. Tight junctions, gap junctions and adherens junctions operate to limit bacterial translocation to the subepithelial layer (Werth *et al.* 2010). Both Crohn's disease and ulcerative colitis are characterized by loss of barrier function and tight junction integrity. Interestingly, dysbiosis of the gut microbiota in IBD is associated with a reduction in Firmicutes (Spor, Koren & Ley 2011, Qin *et al.* 2010). The observation herein that *R. hominis* actively enhances the expression of barrier genes suggests that their loss in IBD patients may be functionally significant. Activation of tight junction complexes is not just the prerogative of *R. hominis*; other commensals, such as *Bacteroides thetaiotaomicron* and *Lactobacillus acidophilus*, also enhance mucosal barrier function (Hooper *et al.* 2001, Ukena *et al.* 2007), inferring probiotic opportunities with these bacteria in human IBD.

The effects of *R. hominis* on the gut immune system are intriguing. The strongest effects were noted in the ascending colon and genes such as *Ly6g6c* were strongly up-regulated, as well as pathways involved in T cell regulation and differentiation and actin polymerization at the immune synapse, which are implicated in T cell activation and effector functions. The most affected T cell pathways included those related to IL-10, ICOS and CTLA-4, which are all involved in supporting Treg differentiation. Importantly, it has been demonstrated that significant increases in CD3⁺CD4⁺CD25⁺FoxP3⁺ cells in the colons of germ-free mice and conventional mice colonized with *R. hominis* using both flow cytometry and immunocytochemistry. These findings complement the recent data on other Clostridium species that drive Treg differentiation. Clearly, *R. hominis* can as a single bacterial species promote mucosal T cell expansion and impact on T cell differentiation.

Flagellin signals are perceived by host TLR5 receptors and many pathogenic flagellin structures induce strong pro-inflammatory responses (Hayashi *et al.* 2001). Signalling through TLR5 in response to resident flagellated commensals may be important for homeostasis, since deletion of TLR5 results in spontaneous colitis in mice (Vijay-Kumar *et al.* 2007). The enhanced expression of *R. hominis* flagellin FlaA1 (RH1) *in vivo* and its potency in activating epithelial cells and BMDCs is of great interest. Other work has shown that *E. coli* flagellin mutants have a colonization advantage over wild-type flagellated strains, possibly due to absence of innate recognition by TLR5 signalling (De Paepe *et al.* 2011, Giraud *et al.* 2008). Herein it is shown that for certain Firmicutes, the expression or possibly up-regulation of flagellin is a natural response to gut colonization. *R. hominis* flagellin protein remains expressed *in vivo* and correlates with sustained colonization, absence of overt inflammation

and expansion of T cells of regulatory phenotype. The recent confirmation of flagellin genes in SFB (Prakash *et al.* 2011, Sczesnak *et al.* 2011) may correlate with the host T cell responses induced by this bacterium (Gaboriau-Routhiau *et al.* 2009, Ivanov *et al.* 2009). Interestingly, RH1 induced a unique effect on both epithelial and DC cultures compared to other flagellins, although all tested structures possess the conserved Arg90 associated with flagellins that bind and activate TLR5 (Yoon *et al.* 2012) suggesting that other sequence/structural differences may account for the unique signalling responses mediated by RH1. The significance of flagellin-TLR5 signalling in Treg responses induced by *R. hominis* was confirmed using TLR5KO. Without wishing to be bound by theory, certain commensal flagellin structures may help to direct immune tolerance responses through TLR5 expressed on either CD103+DC or Treg subsets (Flores-Langarica *et al.* 2012, Kinnebrew *et al.* 2012, Crellin *et al.* 2005). Furthermore, the immune homeostatic effect of *R. hominis* was confirmed in DSS-treated mice, although other signalling moieties, such as butyrate, may also contribute to immune tolerance. This data suggests a potential therapeutic benefit of *R. hominis* in IBD.

An interesting additional biological effect of *R. hominis* colonization was the regulation of genes influencing responses to food and control of appetite. In particular, the satiety hormones *Cck* and *Gcg* were significantly reduced. The effects of *Cck* on food intake are mediated via a vagal afferent pathway. This is the major neural pathway by which information about ingested nutrients reaches the central nervous system to influence both gut function and feeding behaviour. *Cck* acts on the vagal system to decrease expression of molecules that stimulate appetite and feeding, and to increase expression of molecules that inhibit feeding and decrease appetite (*Npy2r* and *Cartpt*, both down-regulated two-fold in the current study). No link between *Cck*, *Gcg* and commensal bacteria has been reported thus far, however, both fatty acids and proteins are potent inducers of *Cck* and *Gcg* (Geraedts *et al.* 2010). *R. hominis* produces short-chain fatty acids such as butyrate with aliphatic tails of less than six carbons; this metabolic activity has been reported to reduce the stimulatory effect on plasma *Cck* observed with longer chain fatty acids (McLaughlin *et al.* 1999). Carcass weight analysis revealed that both body weight and lipid content was indeed significantly increased with *R. hominis*, consistent with body weight increases observed in conventionalization of germfree mice (Turnbaugh *et al.* 2006). Whether this is a direct effect of a reduction in satiety hormones as seen in the current study remains to be seen, as the involvement of *Cck* and *Gcg* has not been reported previously. However, it is important to acknowledge that a link between microbiota colonization and energy harvest from the diet, in part through release of SCFAs, has been shown previously (Turnbaugh *et al.* 2006) (Tremaroli, Kovatcheva-Datchary & Backhed 2010). Given that *R. hominis* is a major butyrate

producer, this mechanism is likely also to contribute to the metabolic efficiency observed following *R. hominis* treatment.

In summary, mono-association of the murine gut with *R. hominis* induced strong bi-directional gene expression events culminating in commensal bacterial adaptation and host tolerance. The flagellin product RH1 seems to exert a unique signaling effect, which preferentially drives expansion of Tregs. The importance of TLR5 in directing Treg differentiation and expansion has been demonstrated. Collectively, this data highlights additional functionality of commensal flagellins, TLR5 signaling and the net direction of the mucosal T cell response.

Functional Assays

In vitro Model

Analysis of the response of intestinal epithelial cells (IECs) to different recombinant flagellins.

Molecular analysis of CCL20 gene expression (a pro-inflammatory gene) after stimulation of IECs to different recombinant flagellins (Figure D.2) pathogenic flagellins (SE, ST, LF, and HM) induced similar but not identical levels of CCL20 mRNA, commensal flagellins displayed much more variable levels of induction. ER (*Eubacterium rectale* 33656), K12 (*Escherichia coli* K12), RH1 and RI3 flagellins induced CCL20 at similar levels to pathogenic flagellins, RI1 and RI4 had intermediate stimulatory activity, RH2 revealed as a low inducer of CCL20 in HT-29 and was devoid of agonistic potential in Caco-2 cells, and RI2 had no observable activity in both cell lines. In conclusion, the inventors distinguished three categories of TLR5 agonists (i) those with no or very low immunostimulatory activity, (ii) those with intermediate immunostimulatory activity and (iii) those with high immunostimulatory activity.

Table D2.

	Unstim	ST	SE	K12	ER	RI1	RI2	RI3	RI4	RH1	RH2	
Caco-2	Unstim	**	**	**	***	*	NS	**	*	*	*	HT-29
	ST	*	**	NS	NS	NS	*	NS	*	NS	*	
	SE	**	NS	*	NS	*	**	NS	*	NS	**	
	K12	**	NS	NS	NS	NS	*	NS	*	NS	*	
	ER	**	NS	NS	NS	*	***	*	**	NS	***	
	RI1	***	NS	*	*	NS	NS	**	NS	*	NS	
	RI2	NS	NS	**	**	***	NS	**	*	*	*	
	RI3	**	NS	*	**	NS	**	NS	**	NS	**	
	RI4	*	NS	**	**	*	NS	***	NS	NS	*	
	RH1	*	NS	NS	NS	NS	*	NS	*	NS	*	
	RH2	NS	NS	**	**	**	NS	**	NS	*	*	

Table D2 indicates significant differences between each treatment calculated with paired t test in HT-29 (upper right) and Caco-2 (lower left). T tests used were unilateral to compare

treatments to unstimulated cells (unstim) and bilateral to compare one treatment with another. NS (non significant); *(p<0.05); ** (p<0.01); *** (p<0.001).

Cellular and immunological analysis of the effects of the recombinant flagellins on the IECs was determined by measurement of cytokines, CXCL8, CXCL10 and CCL2 (MCP-1) secreted. IECs were stimulated for 24hrs with recombinant flagellins (Figure D3).

The flagellins ST, SE, K12, ER, RI3 and RH1 induced the secretion of variable but similar levels of IL-8, IP-10 and MCP-1 chemokines, while RI1, RI2, RI4 and RH2, especially in Caco-2 behaved as low agonists of TLR5, inducing significantly lower amounts of secreted chemokines.

Table D3a

	Unstim	ST	SE	K12	ER	RI1	RI2	RI3	RI4	RH1	RH2
Unstim		*	*	*	*	NS	*	*	*	*	*
ST	NS		NS	NS	NS	NS	NS	NS	NS	NS	NS
SE	*	NS		NS	NS	NS	*	NS	*	*	*
K12	NS	NS	NS		*	NS	**	*	NS	**	*
ER	NS	NS	NS	*		NS	*	**	NS	*	*
RI1	NS	NS	NS	*	*		NS	NS	NS	NS	NS
RI2	NS	NS	**	NS	*	*		NS	**	NS	NS
RI3	NS	NS	NS	NS	NS	NS	NS		NS	*	*
RI4	NS	NS	NS	NS	NS	NS	NS	NS		NS	NS
RH1	*	NS	NS	NS	NS	NS	NS	NS	NS		*
RH2	*	NS	NS	NS	NS	NS	NS	NS	NS	NS	

Table D3b

	Unstim	ST	SE	K12	ER	RI1	RI2	RI3	RI4	RH1	RH2
Unstim		***	***	*	**	NS	NS	*	*	*	NS
ST	*		**	NS	*	***	***	NS	***	NS	*
SE	*	NS		NS	NS	***	***	NS	**	NS	**
K12	*	NS	NS		NS	*	*	NS	*	NS	**
ER	*	NS	NS	NS		**	**	*	**	*	NS
RI1	NS	*	*	*	*		NS	*	**	NS	NS
RI2	NS	*	*	*	*	NS		*	*	NS	NS
RI3	*	NS	NS	NS	NS	*	*		*	NS	*
RI4	*	NS	*	*	*	*	*	*		NS	NS
RH1	NS	NS	NS	NS	NS	NS	NS	NS	NS		NS
RH2	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	

Table D3c

	Unstim	ST	SE	K12	ER	RI1	RI2	RI3	RI4	RH1	RH2
Unstim		NS	*	NS	NS	NS	NS	NS	NS	NS	NS
ST	NS		*	NS	NS	NS	NS	NS	NS	NS	NS
SE	NS	NS		NS	NS	NS	NS	NS	NS	NS	NS
K12	NS	NS	NS		NS	NS	NS	NS	NS	NS	NS
ER	NS	NS	NS	NS		NS	NS	NS	NS	NS	NS
RI1	NS	NS	NS	NS	NS		NS	NS	NS	NS	NS
RI2	NS	NS	NS	NS	NS	NS		NS	NS	NS	NS
RI3	NS	NS	NS	NS	NS	NS	NS		NS	NS	NS
RI4	NS	NS	NS	NS	NS	NS	NS	NS		NS	NS
RH1	NS	NS	NS	NS	NS	NS	NS	NS	NS		NS
RH2	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	

Table D3d

	Unstim	ST	SE	K12	ER	RI1	RI2	RI3	RI4	RH1	RH2
Unstim		*	*	*	**	NS	NS	NS	NS	*	NS
ST	NS		NS	NS	NS	NS	NS	NS	NS	NS	NS
SE	NS	NS		NS	NS	*	NS	*	NS	*	NS
K12	NS	NS	NS		NS	***	**	*	**	NS	**
ER	NS	NS	NS	NS		*	*	NS	*	*	*
RI1	NS	NS	NS	NS	NS		NS	NS	NS	*	NS
RI2	NS	NS	NS	NS	NS	NS		NS	NS	*	NS
RI3	NS	NS	NS	NS	NS	NS	NS		NS	NS	NS
RI4	NS	NS	NS	NS	NS	NS	NS	NS		NS	NS
RH1	NS	NS	NS	NS	NS	NS	NS	NS	NS		**
RH2	NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	

Tables D3a, D3b, D3c and D3d indicate significant differences between each treatment calculated with paired t test. The upper right side of Table D3a and D3b give t values for IL-8 and the lower left side for IP-10, and Tables D3c and D3d give t values for MCP-1. NS (non significant); *(p<0.05); ** (p<0.01); *** (p<0.001).

As shown in Figure D4, neutralization of TLR5 with an anti-TLR5 specific antibody abolished the flagellin-mediated inflammatory response, independent of the commensal or pathogenic origin of flagellin. Therefore, the pro-inflammatory effects of flagellin observed in Caco-2 cells are dependent on TLR5 activation.

Generation of bone marrow-derived dendritic cells and cultures

Bone marrow was harvested from femur and tibia of C3H/HeN and C57Bl6 mice. For GM-CSF-derived dendritic cells, bone marrow cells were resuspended at $1 \times 10^6/\text{mL}$ in RPMI supplemented with 10% FCS and 20ng/mL rmGM-CSF and seeded at 10mL/plate in 100mm² tissue culture plates. After three days culture, loosely adherent cells were collected and replated with GM-CSF supplemented media at $1 \times 10^6/\text{mL}$ in 12 well tissue culture plates. At day 5, cells were stimulated with 100ng/mL flagellins before being harvested on day 6. For Flt3L-derived dendritic cells, bone marrow cells were resuspended at $2 \times 10^6/\text{mL}$ in RPMI supplemented with 10% FCS and 200ng/mL rmFlt3 and seeded at 2mL/well in 12-well tissue culture plates. Cells were cultured for 10 days with Flt3 supplemented media added to each well on day 4. At day 9, cells were stimulated with 100ng/mL flagellins before being harvested on day 10 and analyzed by flow cytometry.

Flow Cytometry Analysis of GM-CSF/IL-4 derived dendritic cells (Figure D5) and Flt3L derived dendritic cells (Figure D6) stimulated with recombinant flagellins was carried out. Flagellin Rh1 was most potent at inducing cellular response in GM-CSF/IL-4 derived dendritic cells with Ri4 and Ri3 having a similar response to commensal flagellins K12 and Er and pathogenic flagellins SE and ST. In contrast Rh2, and Ri2 did not induce cellular response with the GM-CSF/IL-4 derived dendritic cells but significantly increase cellular responses to Flt3L derived dendritic cells. These flagellins and in particular Ri1, are distinctive in their ability to elicit difference in responses with the activation of Flt3L –derived dendritic cells. This response signifies the specificity of the flagellins to a specific subset of dendritic cells. These Flt3L derived dendritic cells are categorised as plasmacytoid dendritic cells which play an important role in immunological tolerance.

In Vivo Model

BOY/J WT and TLR5KO mice were used to evaluate the functional importance of *R. hominis* and its flagellins. The mice were colonised with *R. hominis*. The animals were euthanized and intestinal tissue sampling was performed. Small intestine was collected for immunological analysis by flow cytometry.

Flow cytometry analysis of T cell populations, in particular, T regulatory (Treg) cells, in small intestine lamina propria was carried out (figure 6 A and B). The percentage of FoxP3+CD25+ cells in CD4+ T cell population was significantly higher in the BOY/J WT mice in comparison to the TLR5KO mice. This indicates that *R. hominis* and more specifically the flagellins influences T regs by promoting CD4+ FoxP3+CD25+ regulatory T cells. Therefore it can be concluded that flagellins are important in directing host immune response through TLR5

interactions.

SUMMARY CLAUSES

5 The present invention is defined in the claims and the accompanying description. For convenience other aspects of the present invention are presented herein by way of numbered clauses.

10 1. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or polypeptide FlaA1, and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, for use in regulating the immune system of a subject.

15 2. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or the polypeptide FlaA1, and/or the polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or the polynucleotide encoding said *Roseburia* flagellin, and/or the vector, and/or the host cell comprising said vector, and/or
20 the host cell comprising said polynucleotide sequence, according to clause 1 for use in regulating the adaptive immune system of a subject.

25 3. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or the polypeptide FlaA1, and/or the polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or the polynucleotide encoding said *Roseburia* flagellin, and/or the vector, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, according to clause 1 for use in regulating the innate immune system of a subject.

30 4. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or polypeptide FlaA1, and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or the polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence,
35 for use in maintaining immune homeostasis in a subject.

5. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or polypeptide FlaA1, and/or a polynucleotide sequence

encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, for use in treating an immune disorder in a subject.

6. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or the polypeptide FlaA1, and/or the polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or the polynucleotide encoding said *Roseburia* flagellin, and/or the vector, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, for use according to clause 5, wherein the immune disorder is selected from ulcerative colitis, pouchitis, other autoimmune conditions including rheumatoid arthritis, psoriasis, multiple sclerosis, allergies including coeliac disease, atopic dermatitis and rhinitis.

7. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or polypeptide FlaA1, and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, for use in treating a disorder selected from an inflammatory disorder, an immune disorder and an intestinal disorder in a subject.

8. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or the polypeptide FlaA1, and/or the polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or the polynucleotide encoding said *Roseburia* flagellin, and/or the vector, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, for use according to clause 7, wherein the disorder is selected from irritable bowel syndrome (IBS), colitis, inflammatory bowel disorder (IBD), including Crohn's disease and ulcerative colitis, pouchitis, functional dyspepsia, functional constipation, functional diarrhoea (including antibiotic associated diarrhoea, traveller's diarrhoea and pediatric diarrhoea), functional abdominal pain, functional bloating, Epigastric Pain Syndrome, Postprandial Distress Syndrome, gastrointestinal reflux disease (GERD), autoimmune diseases such as diabetes, arthritis, multiple sclerosis and psoriasis allergies, atopic diseases e.g. atopic dermatitis, necrotising enterocolitis, other infections, and combinations thereof.

9. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species

Roseburia intestinalis), and/or polypeptide FlaA1, and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, for use in improving intestinal microbiota in a subject.

10. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or polypeptide FlaA1, and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, for use in regulating appetite in a subject.

11. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or polypeptide FlaA1, and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, for use in promoting gut health in a subject.

12. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or polypeptide FlaA1, and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, for use in promoting Tregs cells and tolerance mechanisms in the immune system of a subject.

13. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or the polypeptide FlaA1, and/or the polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or the polynucleotide encoding said *Roseburia* flagellin, and/or the vector, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, for use according to any one of clauses 1 to 12, which regulates the induction and/or expression of at least one mobilization or chemotaxis gene.

14. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or the polypeptide FlaA1, and/or the polynucleotide sequence

encoding said polypeptide, and/or *Roseburia* flagellin, and/or the polynucleotide encoding said *Roseburia* flagellin, and/or the vector, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, for use according to clause 13, which upregulates the expression of at least one mobilization or chemotaxis gene, and wherein said gene is selected from *MobA* and *MobL*.

15. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or the polypeptide FlaA1, and/or the polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or the polynucleotide encoding said *Roseburia* flagellin, and/or the vector, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, for use according to any preceding clause, which regulates at least one gene selected from *FlaA1*, *Fla2*, *FlaA3*, and *FlaB*.

16. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or the polypeptide FlaA1, and/or the polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or the polynucleotide encoding said *Roseburia* flagellin, and/or the vector, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, for use according to any preceding clause, which regulates the expression of at least one of the following: acetyl-CoA acetyltransferase, 3-hydroxyacyl-CoA dehydrogenase, butyryl-CoA dehydrogenase, electron transfer flavoprotein beta subunit, electron transfer flavoprotein alpha subunit.

17. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or the polypeptide FlaA1, and/or the polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or the polynucleotide encoding said *Roseburia* flagellin, and/or the vector, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, for use according to any preceding clause, which downregulates the expression of at least one gene selected from *Agt*, *Cartpt*, *Cck*, *Cxcl12* and *Gcg*.

18. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or the polypeptide FlaA1, and/or the polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or the polynucleotide encoding said *Roseburia* flagellin, and/or the vector, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, for use according to any preceding clause, which activates at least one immune response gene in the colon or ileum.

19. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or the polypeptide FlaA1 and/or the polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or the polynucleotide encoding said *Roseburia* flagellin, and/or the vector, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, for use according to any preceding clause, which activates the adaptive immune response by regulating the induction and/or expression of genes associated with T-cell regulation.

20. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or the polypeptide FlaA1, and/or the polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or the polynucleotide encoding said *Roseburia* flagellin, and/or the vector, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, for use according to any preceding clause which upregulates expression of at least one gene selected from *Ly6g6c* and *Ly6g6e* in the ascending colon.

21. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or the polypeptide FlaA1 and/or the polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or the polynucleotide encoding said *Roseburia* flagellin, and/or the vector, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, for use according to any preceding clause which regulates the expression of at least one gene selected from *Tlr5*, *Tlr1*, *Vnn1*, *Defb37*, *Pla2g*, *Muc16*, *Itln*, *Sprr1a*, *Cldn4*, *Pmp22*, *Crb3*, *Magi3*, *Marveld3*, *Mpp7*, *Defcr20*, *Pcgf2*, *Ltbp4*, *Igsf8* and *Tcfe2a*.

22. Use of *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or polypeptide FlaA1, and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, in the preparation of a medicament for regulating the immune system of a subject.

23. Use of *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or the polypeptide FlaA1, and/or the polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or the polynucleotide encoding said *Roseburia* flagellin, and/or the vector, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, according to clause 22

in the preparation of a medicament for regulating the innate immune system of a subject.

24. Use of *Roseburia* (such as a bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or the polypeptide FlaA1, and/or the polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or the polynucleotide encoding said *Roseburia* flagellin, and/or the vector, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, according to clause 22 in the preparation of a medicament for regulating the adaptive immune system of a subject.

25. Use of *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or a polypeptide FlaA1, and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, in the preparation of a medicament for maintaining immune homeostasis in a subject.

26. Use of *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or polypeptide FlaA1 and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, in the preparation of a medicament for treating an immune disorder in a subject.

27. A method of regulating the immune system of a subject, said method comprising administering to the subject a composition comprising *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or polypeptide FlaA1, and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence.

28. A method of activating the innate immune system of a subject, said method comprising administering to the subject a composition comprising *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or polypeptide FlaA1, and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a

vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence.

29. A method of activating the adaptive immune system of a subject, said method comprising administering to the subject a composition comprising *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or polypeptide FlaA1 and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence.

30. A method of treating an immune disorder in a subject, said method comprising administering to the subject a pharmaceutically effective amount of *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or polypeptide FlaA1, and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence.

31. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or the polypeptide FlaA1, and/or the polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or the vector, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence for use according to any one of clauses 1 to 21, or a method according to any one of clauses 27 to 30, or a use according to any one of clauses 22 to 26, wherein the subject is a mammal.

32. *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or polypeptide FlaA1, and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, for use in medicine.

33. A pharmaceutical composition comprising *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or polypeptide FlaA1, and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin,

and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and a pharmaceutically acceptable excipient, carrier or diluent.

34. A nutritional supplement comprising *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or polypeptide FlaA1, and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and a nutritionally acceptable excipient, carrier or diluent.

35. A probiotic composition comprising *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or polypeptide FlaA1, and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence.

36. A feedstuff, food product, dietary supplement, nutritional supplement or food additive comprising *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or polypeptide FlaA1, and/or a polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence.

37. A process for producing a pharmaceutical composition according to clause 32, said process comprising admixing *Roseburia* (such as the bacterial species *Roseburia hominis*, or the bacterial species *Roseburia intestinalis*), and/or the polypeptide FlaA1, and/or the polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or the vector, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, with a pharmaceutically acceptable excipient, carrier or diluent.

38. A process for producing a nutritional supplement according to clause 33, said process comprising admixing *Roseburia* (such as the bacterial species *Roseburia hominis*, or the

bacterial species *Roseburia intestinalis*), and/or the polypeptide FlaA1, and/or the polynucleotide sequence encoding said polypeptide, and/or *Roseburia* flagellin, and/or a polynucleotide encoding said *Roseburia* flagellin, and/or the vector, and/or the host cell comprising said vector, and/or the host cell comprising said polynucleotide sequence, with a nutritionally acceptable excipient, carrier or diluent.

SUMMARY PARAGRAPHS

For convenience other aspects of the present invention are presented herein by way of numbered paragraphs.

1. Polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for use in modulating the inflammation of a tissue or an organ in a subject.

2. The polypeptide or the polynucleotide according to paragraph 1 wherein said polypeptide or polynucleotide reduces the inflammation of the tissue or the organ.

3. The polypeptide or the polynucleotide according to paragraph 2 wherein said polypeptide or polynucleotide reduces the inflammation by epithelial cells of the tissue or the organ.

4. The polypeptide or the polynucleotide according to paragraph 3 wherein said epithelial cells are epithelial cells of the alimentary canal.

5. Polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for use in modulating the production of T cells in a subject; preferably, the polypeptide or the polynucleotide increases the production of T regulatory cells in a subject.

6. Polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for use in restoring immunological tolerance.

7. Polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for use in regulating the immune system and restoring immunological tolerance of a subject.

8. The polypeptide or the polynucleotide according to paragraph 7 for use in regulating the adaptive immune system of a subject.

9. The polypeptide or the polynucleotide according to paragraph 7 for use in regulating the innate immune system of a subject.

10. Polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for use in treating a disorder in a subject, wherein said disorder is an inflammatory disorder and/or an autoimmune disorder.

11. The polypeptide or the polynucleotide according to paragraph 10 wherein said disorder affects the alimentary canal or a section thereof of said subject.

12. The polypeptide or the polynucleotide according to paragraph 10 wherein said disorder is selected from the group consisting of rheumatoid arthritis, psoriasis, multiple sclerosis, type I diabetes, coeliac disease, atopic dermatitis, rhinitis, irritable bowel syndrome (IBS), colitis, inflammatory bowel disorder (IBD), ulcerative colitis, pouchitis, Crohn's disease, functional dyspepsia, atopic diseases, necrotising enterocolitis, and combinations thereof.

13. Polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for use in modulating dendritic cells and/or epithelial cells in a tissue or an organ of a subject.

14. The polypeptide or the polynucleotide according to paragraph 13 wherein said polypeptide or polynucleotide activates dendritic cells and/or epithelial cells.

15. Polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for use in regulating the production of IL-10 and/or TGF β in a cell or cells of a subject.

16. The polypeptide or the polynucleotide according to paragraph 15 wherein the production of IL-10 is by dendritic cells.

17. The polypeptide or the polynucleotide according to paragraph 15 or 16 wherein said polypeptide or polynucleotide upregulates the production of IL-10 and/or TGF β .

18. Polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for use in regulating the production of CD40 and/or I-A/I-E in a cell or cells of a subject.

19. The polypeptide or the polynucleotide according to paragraph 18 wherein the production of CD40 and/or I-A/I-E is by dendritic cells.

20. The polypeptide or the polynucleotide according to paragraph 18 or 19 wherein said polypeptide or polynucleotide upregulates the production CD40 and/or I-A/I-E.

21. Polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for use in regulating the expression of one or more Type I IFN genes in a cell or cells of a subject.
- 5 22. Polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for use in regulating the expression of one or more pro-inflammatory genes in a cell or cells of a subject.
- 10 23. Polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for use in improving intestinal microbiota in a subject.
24. Polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for use in regulating appetite in a subject.
- 15 25. The polypeptide or polynucleotide according to paragraph 24 wherein said polypeptide or polynucleotide stimulates the appetite in the subject.
- 20 26. The polypeptide or polynucleotide according to paragraph 24 or 25 wherein the level of cholecystokinin (Cck) and/or glucagon (Gcg) is reduced in the blood of a subject.
27. Polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for use in regulating the expression of the gene encoding cholecystokinin (Cck) and/or the expression of the gene encoding glucagon (Gcg) in a cell or cells of a subject.
- 25 28. Polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for use in improving alimentary canal health in a subject.
29. The polypeptide or polynucleotide according to any one of paragraphs 1 to 28 wherein said polypeptide or polynucleotide is encapsulated.
- 30 30. A pharmaceutical composition comprising polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide and a pharmaceutically acceptable excipient, carrier or diluent.
- 35 31. The pharmaceutical composition according to paragraph 30 wherein said polypeptide or polynucleotide is encapsulated.

32. A nutritional supplement comprising polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide and a nutritional acceptable excipient, carrier or diluent.

33. The nutritional supplement according to paragraph 32 wherein said polypeptide or polynucleotide is encapsulated.

34. A feedstuff, food product, dietary supplement, or food additive comprising polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide.

35. The feedstuff, food product, dietary supplement, or food additive according to paragraph 34 wherein said feedstuff, food product, dietary supplement, or food additive is encapsulated.

36. A process for producing a pharmaceutical composition according to paragraph 30, said process comprising admixing said polypeptide or polynucleotide with a pharmaceutically acceptable excipient, carrier or diluent; optionally said polypeptide or polynucleotide is encapsulated.

37. A process for producing a nutritional supplement according to paragraph 32, said process comprising admixing said polypeptide or polynucleotide with a nutritionally acceptable excipient, carrier or diluent; optionally said polypeptide or polynucleotide is encapsulated.

38. A method of modulating the inflammation of a tissue or an organ in a subject, said method comprising administering to the subject polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide, and wherein the inflammation of the tissue or organ in the subject is modulated.

39. A method of modulating the production of T cells in a subject, said method comprising administering polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide, and wherein the production of T cells in the subject is modulated, in particular T regulatory cells.

40. A method of regulating the immune system of a subject, said method comprising administering polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide, and wherein the immune system of the subject is regulated.

41. A method of treating a disorder in a subject, said method comprising administering

polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide, wherein said disorder is an inflammatory disorder and/or an autoimmune disorder.

42. A method of modulating dendritic cells and/or epithelial cells in a subject, said method comprising administering polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide, and wherein dendritic cells and/or epithelial cells in the subject are modulated.

43. A method of regulating the production of IL-10 and/or TGF β in a cell or cells of a subject, said method comprising administering polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide to the subject, and wherein the production of IL-10 and/or TGF β in a cell or cells of the subject is regulated.

44. A method of regulating the production of CD40 and/or I-A/I-E in a cell or cells of a subject, said method comprising administering polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide to the subject, and wherein the production of CD40 and/or I-A/I-E in a cell or cells of the subject is regulated.

45. A method of regulating the expression of one of more Type I IFN genes in a cell or cells of a subject, said method comprising administering polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide wherein the expression of one of more Type I IFN genes in a cell or cells of the subject is regulated.

46. A method of regulating the expression of one or more pro-inflammatory genes in a cell or cells of a subject, said method comprising administering polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide, wherein the expression of one or more pro-inflammatory genes in a cell or cells of the subject is regulated.

47. A method of improving intestinal microbiota in a subject, said method comprising administering polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide and wherein the intestinal microbiota in a subject is improved.

48. A method of regulating appetite in a subject, said method comprising administering polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide and wherein the appetite in the subject is regulated.

49. A method of regulating the expression of the gene encoding cholecystokinin (Cck) and/or the expression of the gene encoding glucagon (Gcg) in a cell or cells of a subject,

said method comprising administering polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide and wherein the expression of the gene encoding cholecystokinin (Cck) and/or the expression of the gene encoding glucagon (Gcg) in a cell or cells of the subject is regulated.

50. A method of improving alimentary canal health in a subject, said method comprising administering polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide, and wherein alimentary canal health in a subject is improved.

51. Use of polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for the manufacture of a medicament for modulating the inflammation of a tissue or an organ in a subject.

52. Use of polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for the manufacture of a medicament for modulating the production of T cells in a subject.

53. Use of polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for the manufacture of a medicament for regulating the immune system of a subject.

54. Use of polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for the manufacture of a medicament for the treatment of a disorder in a subject, wherein said disorder is an inflammatory disorder and/or an autoimmune disorder.

55. Use of polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for the manufacture of a medicament for modulating dendritic cells and/or epithelial cells in a tissue or an organ of a subject.

56. Use of polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for the manufacture of a medicament for regulating the production of IL-10 and/or TGF β in a cell or cells of a subject.

57. Use of polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for the manufacture of a medicament for regulating the production of CD40 and/or I-A/I-E in a cell or cells of a subject.

58. Use of polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for the manufacture of a medicament for regulating the expression of one of more Type I IFN

genes in a cell or cells of a subject.

59. Use of polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for the manufacture of a medicament for regulating the expression of one or more pro-inflammatory genes in a cell or cells of a subject.

60. Use of polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for the manufacture of a medicament for improving intestinal microbiota in a subject.

61. Use of polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for the manufacture of a medicament for regulating appetite in a subject.

62. Use of polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for the manufacture of a medicament for regulating the expression of the gene encoding cholecystokinin (Cck) and/or the expression of the gene encoding glucagon (Gcg) in a cell or cells of a subject.

63. Use of polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for the manufacture of a medicament for improving alimentary canal health in a subject.

64. Use of polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide for the manufacture of a medicament for restoring immunological tolerance in a subject.

65. A method of restoring immunological tolerance in a subject, said method comprising administering polypeptide FlaA1 or a polynucleotide sequence encoding said polypeptide and wherein the immunological tolerance in a subject is restored.

REFERENCES

- Aminov, R.I., Walker, A.W., Duncan, S.H., Harmsen, H.J., Welling, G.W. & Flint, H.J. 2006, "Molecular diversity, cultivation, and improved detection by fluorescent in situ hybridization of a dominant group of human gut bacteria related to *Roseburia* spp. or *Eubacterium rectale*", *Applied and Environmental Microbiology*, vol. 72, no. 9, pp. 6371-6376.
- Atarashi, K., Tanoue, T., Shima, T., Imaoka, A., Kuwahara, T., Momose, Y., Cheng, G., Yamasaki, S., Saito, T., Ohba, Y., Taniguchi, T., Takeda, K., Hori, S., Ivanov, I.I., Umesaki, Y., Itoh, K. & Honda, K. 2011, "Induction of colonic regulatory T cells by indigenous *Clostridium* species", *Science (New York, N.Y.)*, vol. 331, no. 6015, pp. 337-341.
- Berg, D.J., Davidson, N., Kuhn, R., Muller, W., Menon, S., Holland, G., Thompson-Snipes, L.,

- Leach, M.W. & Rennick, D. 1996, "Enterocolitis and colon cancer in interleukin-10-deficient mice are associated with aberrant cytokine production and CD4(+) TH1-like responses", *The Journal of clinical investigation*, vol. 98, no. 4, pp. 1010-1020.
- Chung, H. & Kasper, D.L. 2010, "Microbiota-stimulated immune mechanisms to maintain gut homeostasis", *Current opinion in immunology*, vol. 22, no. 4, pp. 455-460.
- Crellin, N.K., Garcia, R.V., Hadisfar, O., Allan, S.E., Steiner, T.S. & Levings, M.K. 2005, "Human CD4+ T cells express TLR5 and its ligand flagellin enhances the suppressive capacity and expression of FOXP3 in CD4+CD25+ T regulatory cells", *Journal of immunology (Baltimore, Md.: 1950)*, vol. 175, no. 12, pp. 8051-8059.
- De Paepe, M., Gaboriau-Routhiau, V., Rainteau, D., Rakotobe, S., Taddei, F. & Cerf-Bensussan, N. 2011, "Trade-off between bile resistance and nutritional competence drives *Escherichia coli* diversification in the mouse gut", *PLoS genetics*, vol. 7, no. 6, pp. e1002107.
- Duck, L.W., Walter, M.R., Novak, J., Kelly, D., Tomasi, M., Cong, Y. & Elson, C.O. 2007, "Isolation of flagellated bacteria implicated in Crohn's disease", *Inflammatory bowel diseases*, vol. 13, no. 10, pp. 1191-1201.
- Duncan, S.H., Aminov, R.I., Scott, K.P., Louis, P., Stanton, T.B. & Flint, H.J. 2006, "Proposal of *Roseburia faecis* sp. nov., *Roseburia hominis* sp. nov. and *Roseburia inulinivorans* sp. nov., based on isolates from human faeces", *International Journal of Systematic and Evolutionary Microbiology*, vol. 56, no. Pt 10, pp. 2437-2441.
- Eckburg, P.B., Bik, E.M., Bernstein, C.N., Purdom, E., Dethlefsen, L., Sargent, M., Gill, S.R., Nelson, K.E. & Relman, D.A. 2005, "Diversity of the human intestinal microbial flora", *Science (New York, N.Y.)*, vol. 308, no. 5728, pp. 1635-1638.
- Flores-Langarica, A., Marshall, J.L., Hitchcock, J., Cook, C., Jobanputra, J., Bobat, S., Ross, E.A., Coughlan, R.E., Henderson, I.R., Uematsu, S., Akira, S. & Cunningham, A.F. 2012, "Systemic flagellin immunization stimulates mucosal CD103+ dendritic cells and drives Foxp3+ regulatory T CELL and IgA responses in the mesenteric lymph node", *Journal of immunology (Baltimore, Md.: 1950)*, vol. 189, no. 12, pp. 5745-5754.
- Frank, D.N., St Amand, A.L., Feldman, R.A., Boedeker, E.C., Harpaz, N. & Pace, N.R. 2007, "Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 34, pp. 13780-13785.
- Gaboriau-Routhiau, V., Rakotobe, S., Lecuyer, E., Mulder, I., Lan, A., Bridonneau, C., Rochet, V., Pisi, A., De Paepe, M., Brandi, G., Eberl, G., Snel, J., Kelly, D. & Cerf-Bensussan, N. 2009, "The key role of segmented filamentous bacteria in the coordinated maturation of gut helper T cell responses", *Immunity*, vol. 31, no. 4, pp. 677-689.
- Geraedts, M.C., Troost, F.J., Tinnemans, R., Soderholm, J.D., Brummer, R.J. & Saris, W.H. 2010, "Release of satiety hormones in response to specific dietary proteins is different

between human and murine small intestinal mucosa", *Annals of Nutrition & Metabolism*, vol. 56, no. 4, pp. 308-313.

Geuking, M.B., Cahenzli, J., Lawson, M.A., Ng, D.C., Slack, E., Hapfelmeier, S., McCoy, K.D. & Macpherson, A.J. 2011, "Intestinal bacterial colonization induces mutualistic regulatory T cell responses", *Immunity*, vol. 34, no. 5, pp. 794-806.

Giraud, A., Arous, S., De Paepe, M., Gaboriau-Routhiau, V., Bambou, J.C., Rakotobe, S., Lindner, A.B., Taddei, F. & Cerf-Bensussan, N. 2008, "Dissecting the genetic components of adaptation of *Escherichia coli* to the mouse gut", *PLoS genetics*, vol. 4, no. 1, pp. e2.

Hapfelmeier, S., Lawson, M.A., Slack, E., Kirundi, J.K., Stoel, M., Heikenwalder, M., Cahenzli, J., Velykoredko, Y., Balmer, M.L., Endt, K., Geuking, M.B., Curtiss, R., 3rd, McCoy, K.D. & Macpherson, A.J. 2010, "Reversible microbial colonization of germ-free mice reveals the dynamics of IgA immune responses", *Science (New York, N.Y.)*, vol. 328, no. 5986, pp. 1705-1709.

Hayashi, F., Smith, K.D., Ozinsky, A., Hawn, T.R., Yi, E.C., Goodlett, D.R., Eng, J.K., Akira, S., Underhill, D.M. & Aderem, A. 2001, "The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5", *Nature*, vol. 410, no. 6832, pp. 1099-1103.

Hooper, L.V., Wong, M.H., Thelin, A., Hansson, L., Falk, P.G. & Gordon, J.I. 2001, "Molecular analysis of commensal host-microbial relationships in the intestine", *Science (New York, N.Y.)*, vol. 291, no. 5505, pp. 881-884.

Hossain, M.S., Jaye, D.L., Pollack, B.P., Farris, A.B., Tselanyane, M.L., David, E., Roback, J.D., Gewirtz, A.T. & Waller, E.K. 2011, "Flagellin, a TLR5 agonist, reduces graft-versus-host disease in allogeneic hematopoietic stem cell transplantation recipients while enhancing antiviral immunity", *Journal of immunology (Baltimore, Md.: 1950)*, vol. 187, no. 10, pp. 5130-5140.

Ivanov, I.I., Atarashi, K., Manel, N., Brodie, E.L., Shima, T., Karaoz, U., Wei, D., Goldfarb, K.C., Santee, C.A., Lynch, S.V., Tanoue, T., Imaoka, A., Itoh, K., Takeda, K., Umesaki, Y., Honda, K. & Littman, D.R. 2009, "Induction of intestinal Th17 cells by segmented filamentous bacteria", *Cell*, vol. 139, no. 3, pp. 485-498.

Kang, S., Denman, S.E., Morrison, M., Yu, Z., Dore, J., Leclerc, M. & McSweeney, C.S.

2010, "Dysbiosis of fecal microbiota in Crohn's disease patients as revealed by a custom phylogenetic microarray", *Inflammatory bowel diseases*, vol. 16, no. 12, pp. 2034-2042.

Kelly, D., Campbell, J.I., King, T.P., Grant, G., Jansson, E.A., Coutts, A.G., Pettersson, S. & Conway, S. 2004, "Commensal anaerobic gut bacteria attenuate inflammation by regulating nuclear-cytoplasmic shuttling of PPAR-gamma and RelA", *Nature immunology*, vol. 5, no. 1, pp. 104-112.

Kinnebrew, M.A., Buffie, C.G., Diehl, G.E., Zenewicz, L.A., Leiner, I., Hohl, T.M., Flavell, R.A., Littman, D.R. & Pamer, E.G. 2012, "Interleukin 23 production by intestinal

CD103(+)CD11b(+) dendritic cells in response to bacterial flagellin enhances mucosal innate immune defense", *Immunity*, vol. 36, no. 2, pp. 276-287.

Letran, S.E., Lee, S.J., Atif, S.M., Flores-Langarica, A., Uematsu, S., Akira, S., Cunningham, A.F. & McSorley, S.J. 2011, "TLR5-deficient mice lack basal inflammatory and metabolic defects but exhibit impaired CD4 T cell responses to a flagellated pathogen", *Journal of immunology (Baltimore, Md.: 1950)*, vol. 186, no. 9, pp. 5406-5412.

Machiels K., Joossens M., Sabino J., De Preter V., Arijis I., Ballet V., Claes K., Verhaegen J., Van Assche G., Rutgeerts P. & Vermeire S. 2013, "Predominant dysbiosis in patients with ulcerative colitis is different from Crohn's disease patients", *Inflammatory Bowel Diseases. 8th Congress of ECCO*, Feb 14-16, 2013.

Macpherson, A.J. 2006, "IgA adaptation to the presence of commensal bacteria in the intestine", *Current topics in microbiology and immunology*, vol. 308, pp. 117-136.

Macpherson, A.J., Hunziker, L., McCoy, K. & Lamarre, A. 2001, "IgA responses in the intestinal mucosa against pathogenic and non-pathogenic microorganisms", *Microbes and infection / Institut Pasteur*, vol. 3, no. 12, pp. 1021-1035.

Macpherson, A.J., Martinic, M.M. & Harris, N. 2002, "The functions of mucosal T cells in containing the indigenous commensal flora of the intestine", *Cellular and molecular life sciences : CMLS*, vol. 59, no. 12, pp. 2088-2096.

Mahowald, M.A., Rey, F.E., Seedorf, H., Turnbaugh, P.J., Fulton, R.S., Wollam, A., Shah, N., Wang, C., Magrini, V., Wilson, R.K., Cantarel, B.L., Coutinho, P.M., Henrissat, B., Crock, L.W., Russell, A., Verberkmoes, N.C., Hettich, R.L. & Gordon, J.I. 2009, "Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla", *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 14, pp. 5859-5864.

Mallya, M., Campbell, R.D. & Aguado, B. 2006, "Characterization of the five novel Ly-6 superfamily members encoded in the MHC, and detection of cells expressing their potential ligands", *Protein science : a publication of the Protein Society*, vol. 15, no. 10, pp. 2244-2256.

Mazmanian, S.K., Liu, C.H., Tzianabos, A.O. & Kasper, D.L. 2005, "An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system", *Cell*, vol. 122, no. 1, pp. 107-118.

McLaughlin, J., Grazia Luca, M., Jones, M.N., D'Amato, M., Dockray, G.J. & Thompson, D.G. 1999, "Fatty acid chain length determines cholecystokinin secretion and effect on human gastric motility", *Gastroenterology*, vol. 116, no. 1, pp. 46-53.

Monteleone, I., Platt, A.M., Jaensson, E., Agace, W.W. & Mowat, A.M. 2008, "IL-10-dependent partial refractoriness to Toll-like receptor stimulation modulates gut mucosal dendritic cell function", *European journal of immunology*, vol. 38, no. 6, pp. 1533-1547.

- Nutsch, K.M. & Hsieh, C.S. 2012, "T cell tolerance and immunity to commensal bacteria", *Current opinion in immunology*, vol. 24, no. 4, pp. 385-391.
- Prakash, T., Oshima, K., Morita, H., Fukuda, S., Imaoka, A., Kumar, N., Sharma, V.K., Kim, S.W., Takahashi, M., Saitou, N., Taylor, T.D., Ohno, H., Umesaki, Y. & Hattori, M. 2011, "Complete genome sequences of rat and mouse segmented filamentous bacteria, a potent inducer of th17 cell differentiation", *Cell host & microbe*, vol. 10, no. 3, pp. 273-284.
- Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., Mende, D.R., Li, J., Xu, J., Li, S., Li, D., Cao, J., Wang, B., Liang, H., Zheng, H., Xie, Y., Tap, J., Lepage, P., Bertalan, M., Balto, J.M., Hansen, T., Le Paslier, D., Linneberg, A., Nielsen, H.B., Pelletier, E., Renault, P., Sicheritz-Ponten, T., Turner, K., Zhu, H., Yu, C., Li, S., Jian, M., Zhou, Y., Li, Y., Zhang, X., Li, S., Qin, N., Yang, H., Wang, J., Brunak, S., Dore, J., Guarner, F., Kristiansen, K., Pedersen, O., Parkhill, J., Weissenbach, J., MetaHIT Consortium, Bork, P., Ehrlich, S.D. & Wang, J. 2010, "A human gut microbial gene catalogue established by metagenomic sequencing", *Nature*, vol. 464, no. 7285, pp. 59-65.
- Round, J.L., Lee, S.M., Li, J., Tran, G., Jabri, B., Chatila, T.A. & Mazmanian, S.K. 2011, "The Toll-like receptor 2 pathway establishes colonization by a commensal of the human microbiota", *Science (New York, N.Y.)*, vol. 332, no. 6032, pp. 974-977.
- Scanlan, P.D., Shanahan, F., O'Mahony, C. & Marchesi, J.R. 2006, "Culture-independent analyses of temporal variation of the dominant fecal microbiota and targeted bacterial subgroups in Crohn's disease", *Journal of clinical microbiology*, vol. 44, no. 11, pp. 3980-3988.
- Schulke, S., Burggraf, M., Waibler, Z., Wangorsch, A., Wolfheimer, S., Kalinke, U., Vieths, S., Toda, M. & Scheurer, S. 2011, "A fusion protein of flagellin and ovalbumin suppresses the TH2 response and prevents murine intestinal allergy", *The Journal of allergy and clinical immunology*, vol. 128, no. 6, pp. 1340-1348.e12.
- Sczesnak, A., Segata, N., Qin, X., Gevers, D., Petrosino, J.F., Huttenhower, C., Littman, D.R. & Ivanov, I.I. 2011, "The genome of th17 cell-inducing segmented filamentous bacteria reveals extensive auxotrophy and adaptations to the intestinal environment", *Cell host & microbe*, vol. 10, no. 3, pp. 260-272.
- Spor, A., Koren, O. & Ley, R. 2011, "Unravelling the effects of the environment and host genotype on the gut microbiome", *Nature reviews.Microbiology*, vol. 9, no. 4, pp. 279-290.
- Tremaroli, V., Kovatcheva-Datchary, P. & Backhed, F. 2010, "A role for the gut microbiota in energy harvesting?", *Gut*, vol. 59, no. 12, pp. 1589-1590.
- Turnbaugh, P.J., Backhed, F., Fulton, L. & Gordon, J.I. 2008, "Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome", *Cell host & microbe*, vol. 3, no. 4, pp. 213-223.

- Turnbaugh, P.J., Ley, R.E., Mahowald, M.A., Magrini, V., Mardis, E.R. & Gordon, J.I. 2006, "An obesity-associated gut microbiome with increased capacity for energy harvest", *Nature*, vol. 444, no. 7122, pp. 1027-1031.
- Ukena, S.N., Singh, A., Dringenberg, U., Engelhardt, R., Seidler, U., Hansen, W., Bleich, A., Bruder, D., Franzke, A., Rogler, G., Suerbaum, S., Buer, J., Gunzer, F. & Westendorf, A.M. 2007, "Probiotic *Escherichia coli* Nissle 1917 inhibits leaky gut by enhancing mucosal integrity", *PloS one*, vol. 2, no. 12, pp. e1308.
- Vijay-Kumar, M., Sanders, C.J., Taylor, R.T., Kumar, A., Aitken, J.D., Sitaraman, S.V., Neish, A.S., Uematsu, S., Akira, S., Williams, I.R. & Gewirtz, A.T. 2007, "Deletion of TLR5 results in spontaneous colitis in mice", *The Journal of clinical investigation*, vol. 117, no. 12, pp. 3909-3921.
- Werth, M., Walentin, K., Aue, A., Schonheit, J., Wuebken, A., Pode-Shakked, N., Vilianovitch, L., Erdmann, B., Dekel, B., Bader, M., Barasch, J., Rosenbauer, F., Luft, F.C. & Schmidt-Ott, K.M. 2010, "The transcription factor grainyhead-like 2 regulates the molecular composition of the epithelial apical junctional complex", *Development (Cambridge, England)*, vol. 137, no. 22, pp. 3835-3845.
- Wilson, R.H., Maruoka, S., Whitehead, G.S., Foley, J.F., Flake, G.P., Sever, M.L., Zeldin, D.C., Kraft, M., Garantziotis, S., Nakano, H. & Cook, D.N. 2012, "The Toll-like receptor 5 ligand flagellin promotes asthma by priming allergic responses to indoor allergens", *Nature medicine*, vol. 18, no. 11, pp. 1705-1710.
- Yoon, S.I., Kurnasov, O., Natarajan, V., Hong, M., Gudkov, A.V., Osterman, A.L. & Wilson, I.A. 2012, "Structural basis of TLR5-flagellin recognition and signaling", *Science (New York, N.Y.)*, vol. 335, no. 6070, pp. 859-864.

Reference List for Supplementary Information

- (1) Genome sequence assembly using trace signals and additional sequence information. Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB); 1999.
- (2) Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, *et al*. The RAST Server: rapid annotations using subsystems technology. BMC Genomics 2008 Feb 8;9:75-2164-9-75.
- (3) Dennis G,Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, *et al*. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 2003;4(5):P3.
- (4) Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JA. Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res 2007 Jul;35(Web Server issue):W71-4.
- (5) Duck LW, Walter MR, Novak J, Kelly D, Tomasi M, Cong Y, *et al*. Isolation of flagellated bacteria implicated in Crohn's disease. Inflamm Bowel Dis 2007 Oct;13(10):1191-1201.

- (6) Berg DJ, Davidson N, Kuhn R, Muller W, Menon S, Holland G, Thompson-Snipes L, Leach MW, Rennick D. Enterocolitis and colon cancer in interleukin-10-deficient mice are associated with aberrant cytokine production and CD4(+) TH1-like responses. *J.Clin.Invest.*, 1996, 98, 4, 1010-1020.
- 5 (7) Monteleone I, Platt AM, Jaensson E, Agace WW, Mowat AM. IL-10-dependent partial refractoriness to Toll-like receptor stimulation modulates gut mucosal dendritic cell function. *Eur J Immunol.* 2008 38(6):1533-47
- (8) Inaba K, Inaba M, Romani N, Aya H, Deguchi M, Ikehara S, *et al.* Generation of large numbers of dendritic cells from mouse bone marrow cultures supplemented with
- 10 granulocyte/macrophage colony-stimulating factor. *J Exp Med* 1992 Dec 1;176(6):1693-1702.
- (9) Brasel K, De Smedt T, Smith JL, Maliszewski CR. Generation of murine dendritic cells from flt3-ligand-supplemented bone marrow cultures. *Blood* 2000 Nov 1;96(9):3029-3039.
- (10) Weigel BJ, Nath N, Taylor PA, Panoskaltsis-Mortari A, Chen W, Krieg AM, *et al.*
- 15 Comparative analysis of murine marrow-derived dendritic cells generated by Flt3L or GM-CSF/IL-4 and matured with immune stimulatory agents on the in vivo induction of antileukemia responses. *Blood* 2002 Dec 1;100(12):4169-4176.
- (10) Xu Y, Zhan Y, Lew AM, Naik SH, Kershaw MH. Differential development of murine dendritic cells by GM-CSF versus Flt3 ligand has implications for inflammation and
- 20 trafficking. *J Immunol* 2007 Dec 1;179(11):7577-7584.
- (12) Olivera L, Canul RR, Pereira-Pacheco F, Cockburn J, Soldani F, McKenzie NH, *et al.* Nutritional and physiological responses of young growing rats to diets containing raw cowpea seed meal, protein isolate (globulins), or starch. *J Agric Food Chem* 2003 Jan 1;51(1):319-325.
- 25 All publications mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described methods and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly
- 30 limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in biochemistry and molecular biology or related fields are intended to be within the scope of the following claims.

CLAIMS

1. A method of treating a disorder in a subject, said method comprising administering to the subject a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, wherein said disorder is an inflammatory disorder and/or an autoimmune disorder, and wherein *Roseburia hominis* bacteria *per se* are excluded.

2. Use of a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, for the manufacture of a medicament for the treatment of a disorder in a subject, wherein said disorder is an inflammatory disorder and/or an autoimmune disorder, and wherein *Roseburia hominis* bacteria *per se* are excluded.

3. The method of claim 1 or the use of claim 2, wherein said *Roseburia* flagellin is selected from the group consisting of:

polypeptides having at least 95% identity to SEQ ID NO 2, SEQ ID NO 4, SEQ ID NO 6, SEQ ID NO 8, SEQ ID NO 10 or SEQ ID NO 12 or variants, homologues, fragments or derivatives thereof and combinations thereof;

and wherein the polynucleotide sequence encoding said *Roseburia* flagellin is selected from the group consisting of:

polynucleotide sequences encoding a polypeptide having at least 95% identity to SEQ ID NO 2, SEQ ID NO 4, SEQ ID NO 6, SEQ ID NO 8, SEQ ID NO 10 or SEQ ID NO 12 or variants, homologues, fragments or derivatives thereof;

polynucleotide sequences having at least 95% identity to SEQ ID NO 1, SEQ ID NO 3, SEQ ID NO 5, SEQ ID NO 7, SEQ ID NO 9 or SEQ ID NO 11 or variants, homologues, fragments or derivatives thereof;

and combinations thereof.

4. The method or use according to any one of claims 1 to 3, wherein the disorder is selected from the group consisting of rheumatoid arthritis, psoriasis, multiple sclerosis, type I diabetes, coeliac disease, atopic dermatitis, rhinitis, irritable bowel syndrome (IBS), colitis, inflammatory bowel disorder (IBD), ulcerative colitis, pouchitis, Crohn's disease, functional dyspepsia, atopic diseases, necrotising enterocolitis, and combinations thereof.

5. The method or use of any one of claims 1 to 3, wherein the *Roseburia* flagellin, or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, increases the production of T regulatory cells in a subject.

6. The method or use of any one of claims 1 to 5, wherein the *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, increases the production of IL-10 and/or TGF β in a cell or cells of a subject.

7. The method or use of any one of claims 1 to 6, wherein the *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, increases the production of cell surface markers involved in immune responses and antigen recognition such as CD40, I-A/I-E, CD317/BST-2, CD103, CD80, CD86, CD83 and/or Siglec-H and/or the species equivalent in a cell or cells of a subject.

8. The method or use of claim 7, wherein the production of cell surface markers involved in immune responses and antigen recognition such as CD40, I-A/I-E, CD317/BST-2, CD103, CD80, CD86, CD83 and/or Siglec-H and/or the species equivalent is by dendritic cells.

9. The method or use of any one of claims 1 to 8, wherein the *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, decreases the expression of one or more Type I IFN genes in a cell or cells of a subject.

10. A method of stimulating the appetite in a subject, said method comprising administering to the subject a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and wherein the appetite in the subject is stimulated, and

wherein *Roseburia hominis* bacteria *per se* are excluded.

11. Use of a *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, for the manufacture of a medicament for stimulating the appetite in a subject, and wherein *Roseburia hominis* bacteria *per se* are excluded.

12. The method or use of any one of claims 10 to 11 wherein, the level of cholecystokinin (Cck) and/or glucagon (Gcg) is reduced in the blood of a subject.

13. The method or use of any one of claims 1 to 12, wherein the host cell is bacteria.

14. A pharmaceutical composition and/or a nutritional supplement comprising *Roseburia* flagellin, and/or a polynucleotide sequence encoding said *Roseburia* flagellin, and/or a vector comprising said polynucleotide sequence, and/or a host cell comprising said vector, and/or a host cell comprising said polynucleotide sequence, and a pharmaceutically and/or nutritionally acceptable excipient, carrier or diluent, and wherein a pharmaceutical composition and/or nutritional supplement comprising *Roseburia hominis* bacteria *per se* is excluded.

15. The pharmaceutical composition and/or nutritional supplement according to claim 14 wherein said *Roseburia* flagellin, and/or said polynucleotide sequence, and/or said vector, and/or said host cell comprising said vector, and/or said host cell comprising said polynucleotide sequence, is encapsulated.

16. The pharmaceutical composition and/or nutritional supplement according to claim 14 or claim 15, wherein the host cell is bacteria.

17. A method of treating a disorder in a subject, said method comprising administering to the subject the pharmaceutical composition and/or nutritional supplement of any one of claims 14 to 16, wherein said disorder is an inflammatory disorder and/or an autoimmune disorder.

18. Use of the pharmaceutical composition and/or nutritional supplement of any one of claims 14 to 16, for the manufacture of a medicament for the treatment of a disorder in a subject, wherein said disorder is an inflammatory disorder and/or an autoimmune disorder.

19. A method of stimulating the appetite in a subject, said method comprising administering to the subject the pharmaceutical composition and/or nutritional supplement of any one of claims 14 to 16.

20. Use of the pharmaceutical composition and/or nutritional supplement of any one of claims 14 to 16, for the manufacture of a medicament for stimulating the appetite in a subject.

21. A process for producing a pharmaceutical composition and/or nutritional supplement according to claim 15, said process comprising admixing said *Roseburia* flagellin, and/or said polynucleotide sequence, and/or said vector, and/or said host cell comprising said vector, and/or said host cell comprising said polynucleotide sequence, with a pharmaceutically or nutritionally acceptable excipient, carrier or diluent; optionally said *Roseburia* flagellin, and/or said polynucleotide sequence, and/or said vector, and/or said host cell comprising said vector, and/or said host cell comprising said polynucleotide sequence, is encapsulated.

22. The process according to claim 21, wherein the host cell is bacteria.

23. The method or use of any one of claims 1 to 13 and 17 to 20, the pharmaceutical composition and/or nutritional supplement of any one of claims 14 to 16, and the process of claim 21 or claim 22, wherein the *Roseburia* flagellin is derived or derivable from *Roseburia hominis*.

24. The method, use, pharmaceutical composition, nutritional supplement or process of claim 23, wherein the *Roseburia hominis* flagellin is a polypeptide having at least 99% identity to SEQ ID NO 2 or SEQ ID NO 4, and wherein the polynucleotide sequence encoding said *Roseburia hominis* flagellin is a polynucleotide sequence encoding a polypeptide having at least 99% identity to SEQ ID NO 2 or SEQ ID NO 4, or a polynucleotide sequence having at least 99% identity to SEQ ID NO 1 or SEQ ID NO 3.

25. The method, use, pharmaceutical composition, nutritional supplement or process of claim 23 or claim 24, wherein the *Roseburia hominis* flagellin has the polypeptide sequence shown as SEQ ID NO 2 or SEQ ID NO 4, and wherein the polynucleotide sequence encoding said *Roseburia hominis* flagellin has the polynucleotide sequence encoding a polypeptide which has the sequence shown as SEQ ID NO 2 or SEQ ID NO 4, or the polynucleotide sequence shown as SEQ ID NO 1 or SEQ ID NO 3.

Figure 1

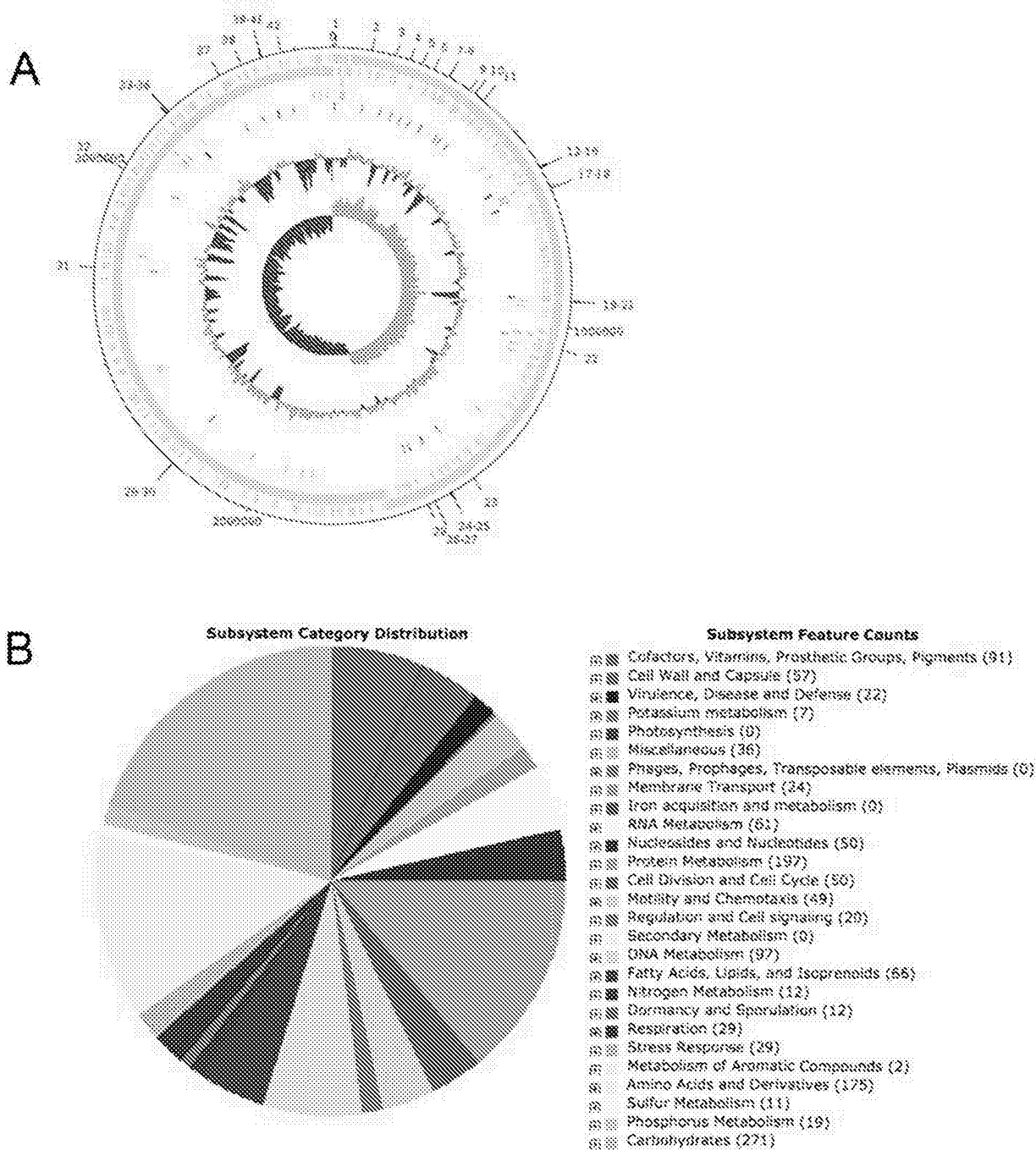


Figure 2

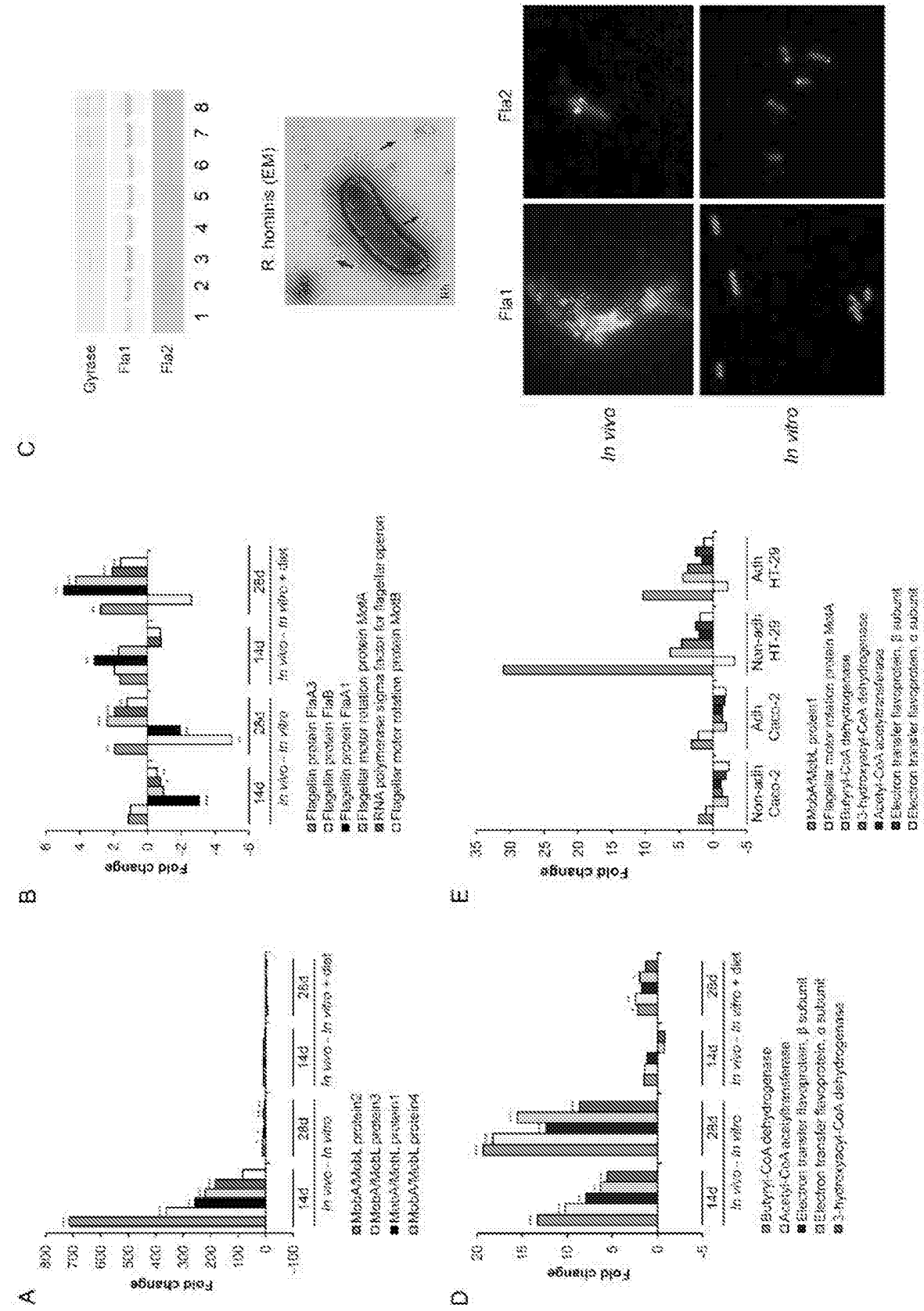


Figure 3

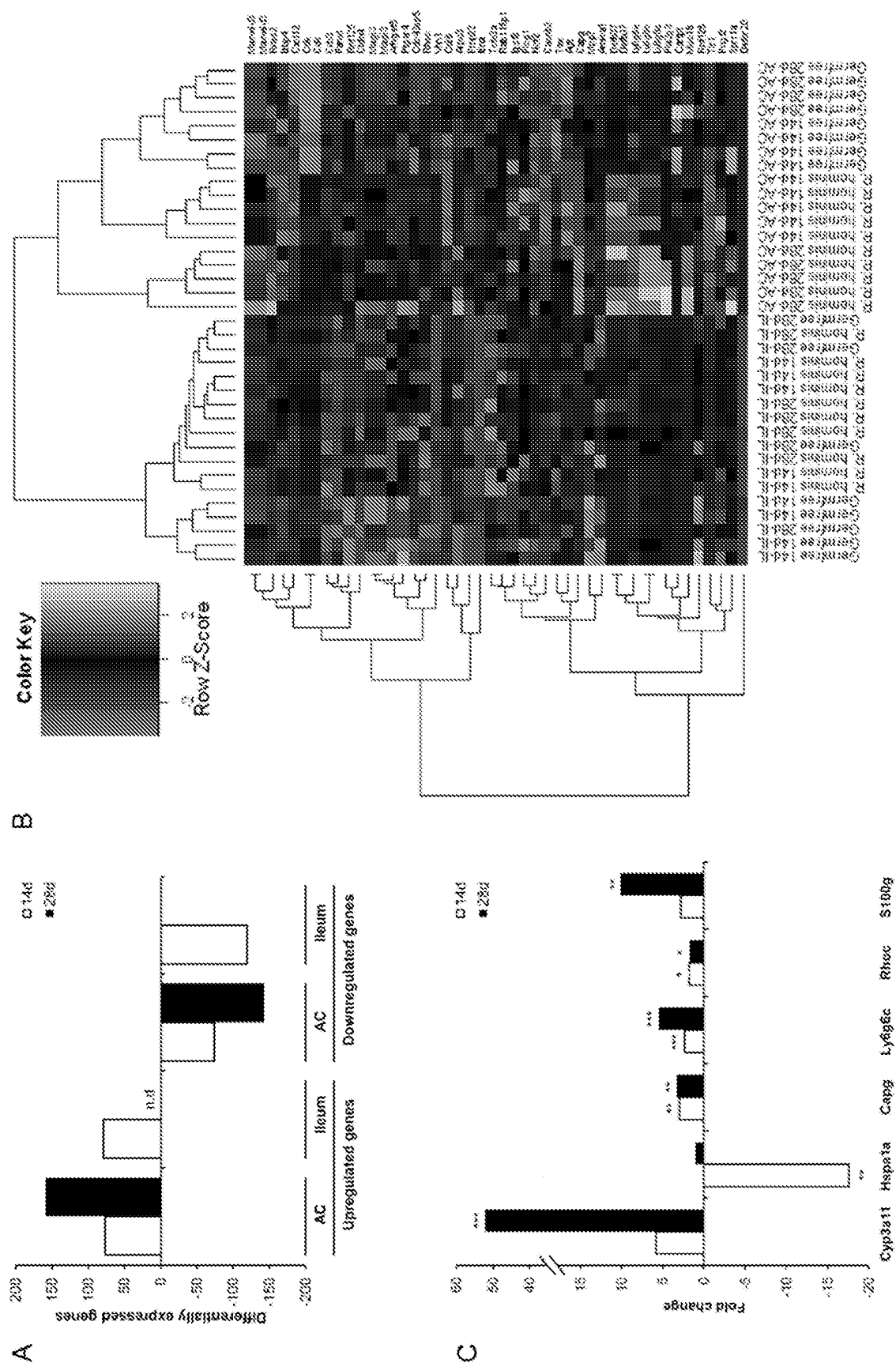


Figure 4

Gated on live, singlets, lymphocytes, B220- CD3+ CD8- CD4+ subset

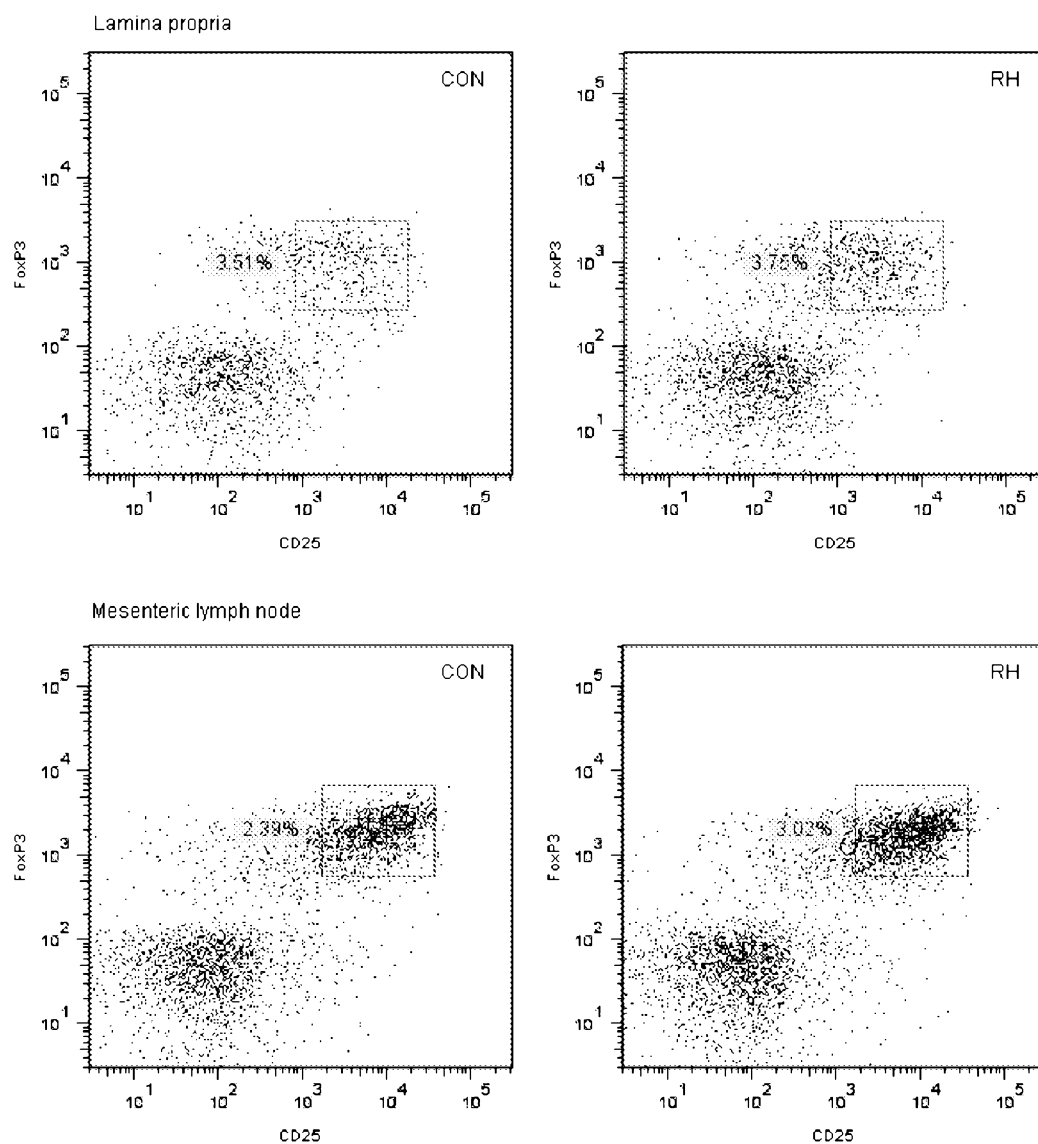


Figure 5

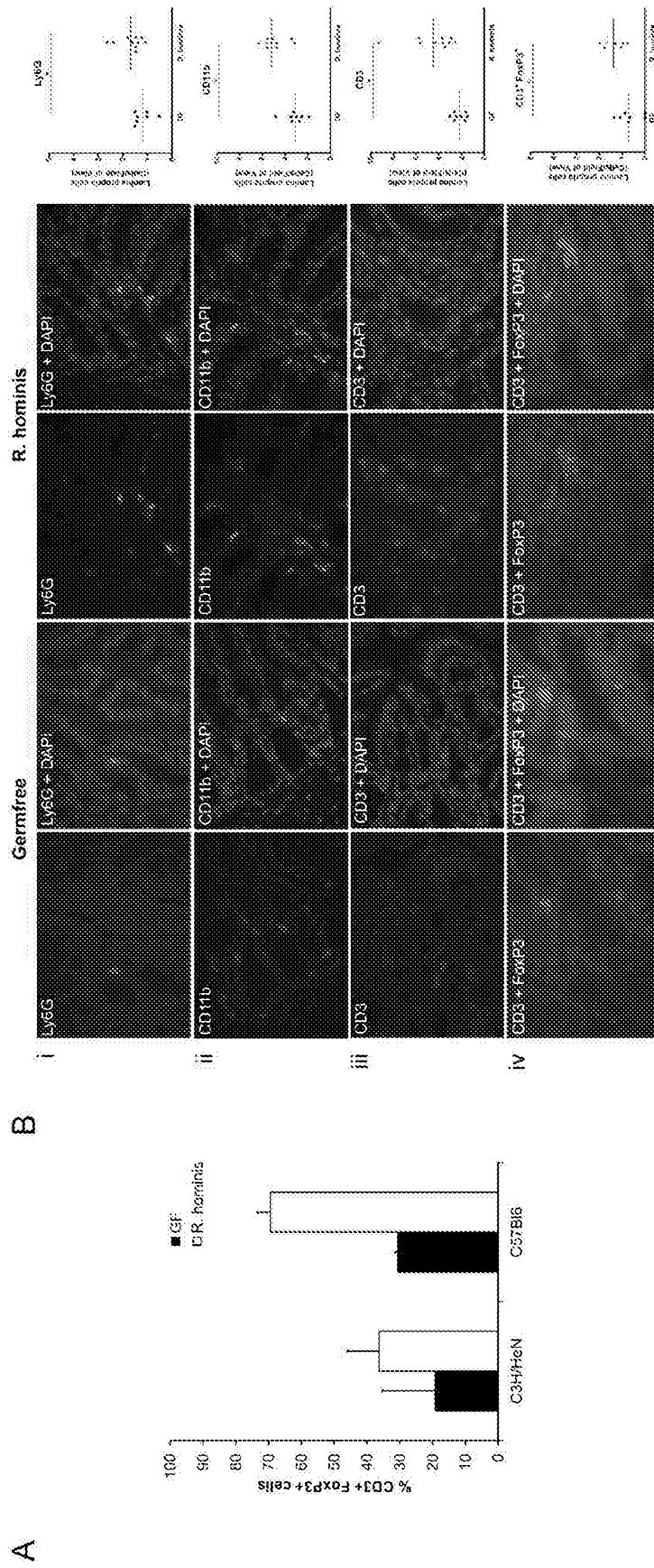


Figure 6

Gated on live, singlet, CD11c⁺B220⁺CD317⁺ cells

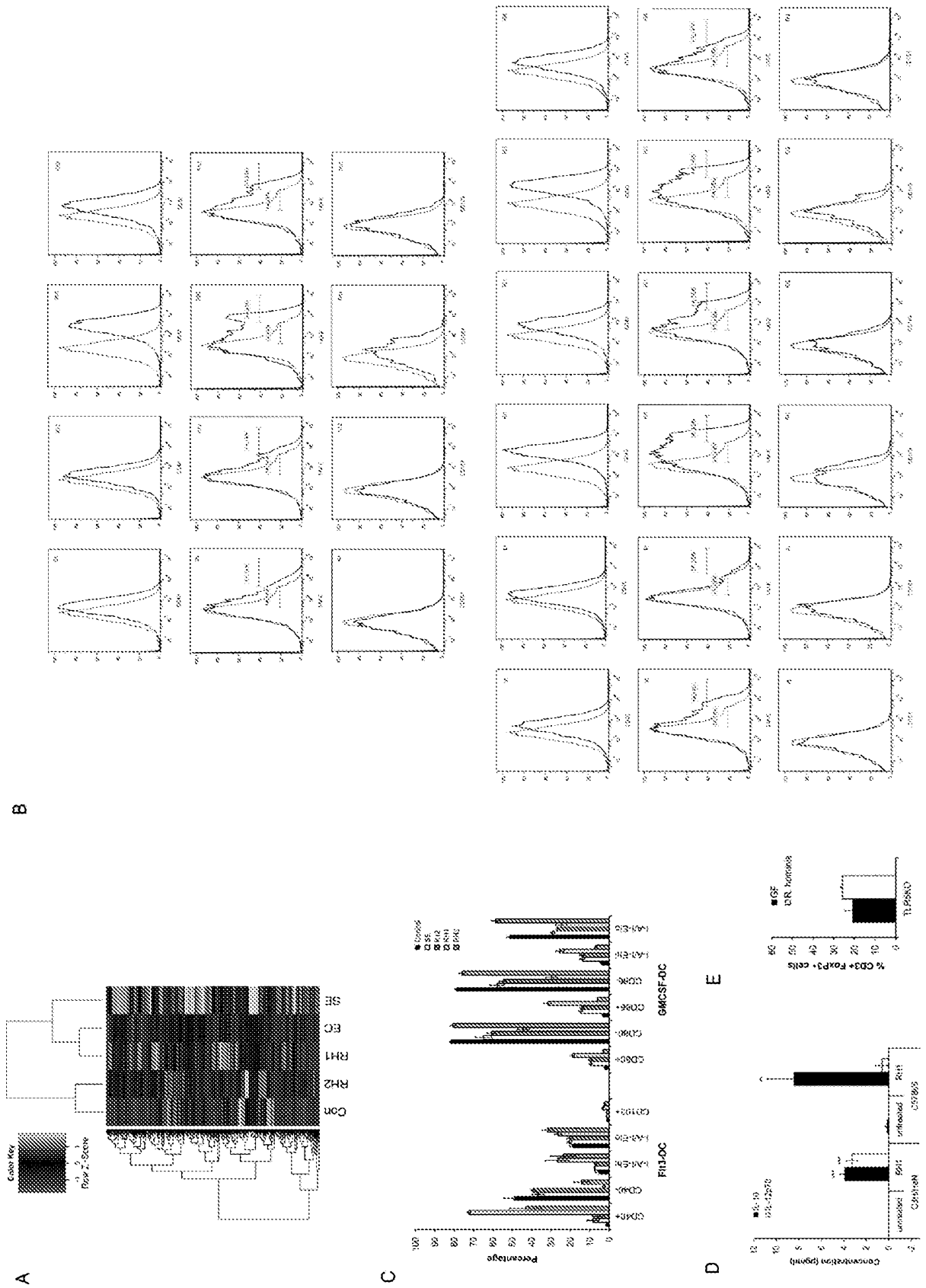


Figure 7

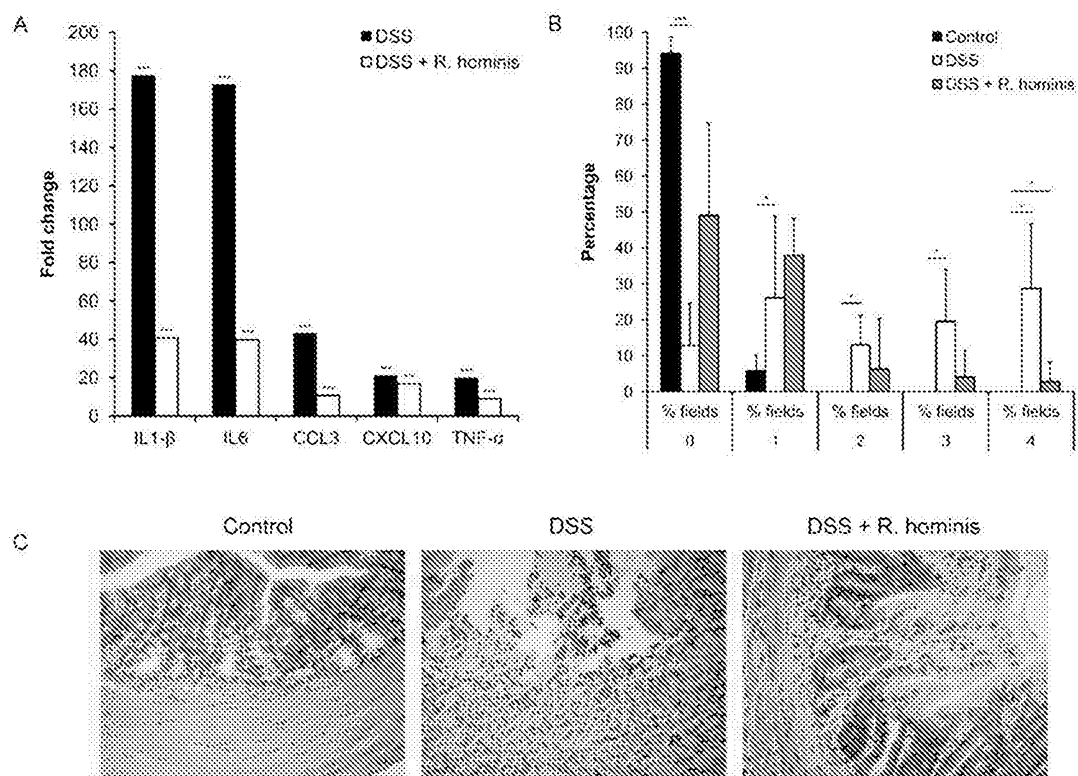


Figure S1

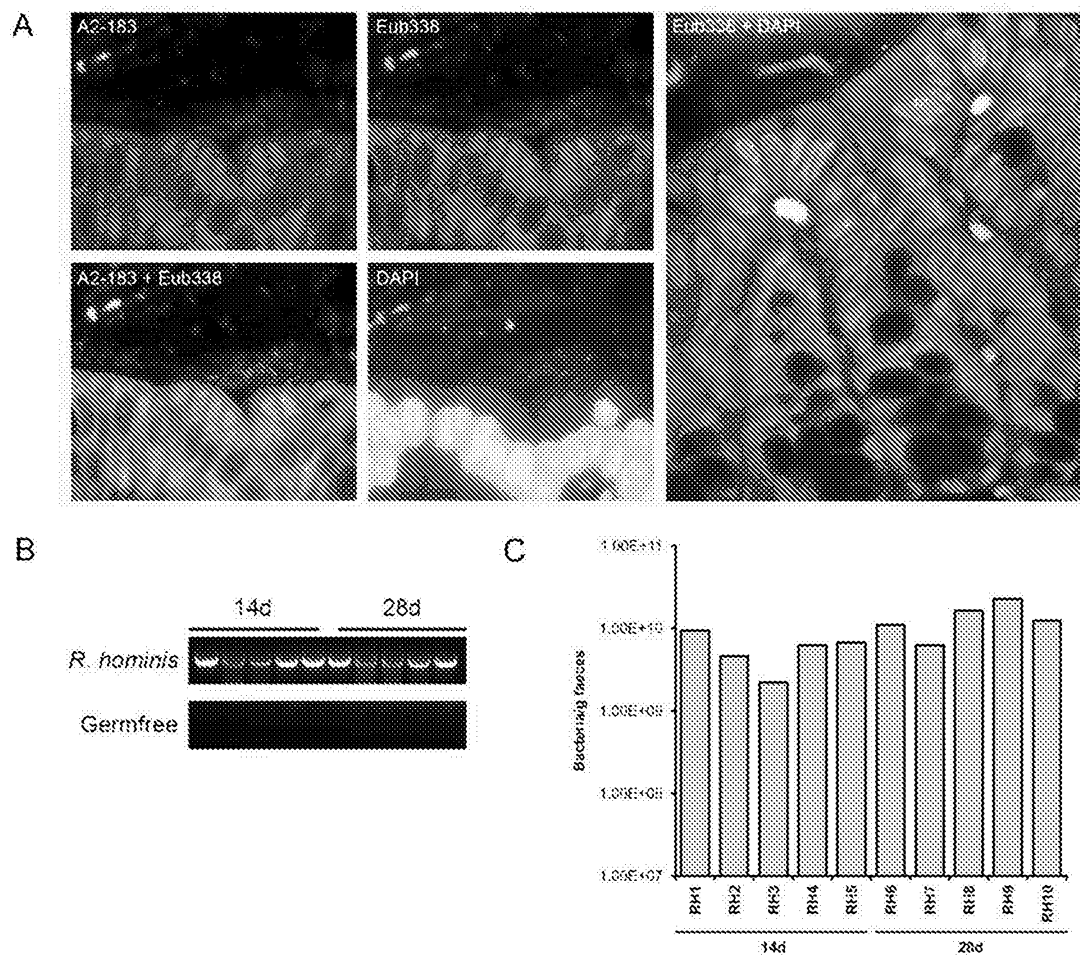


Figure S2

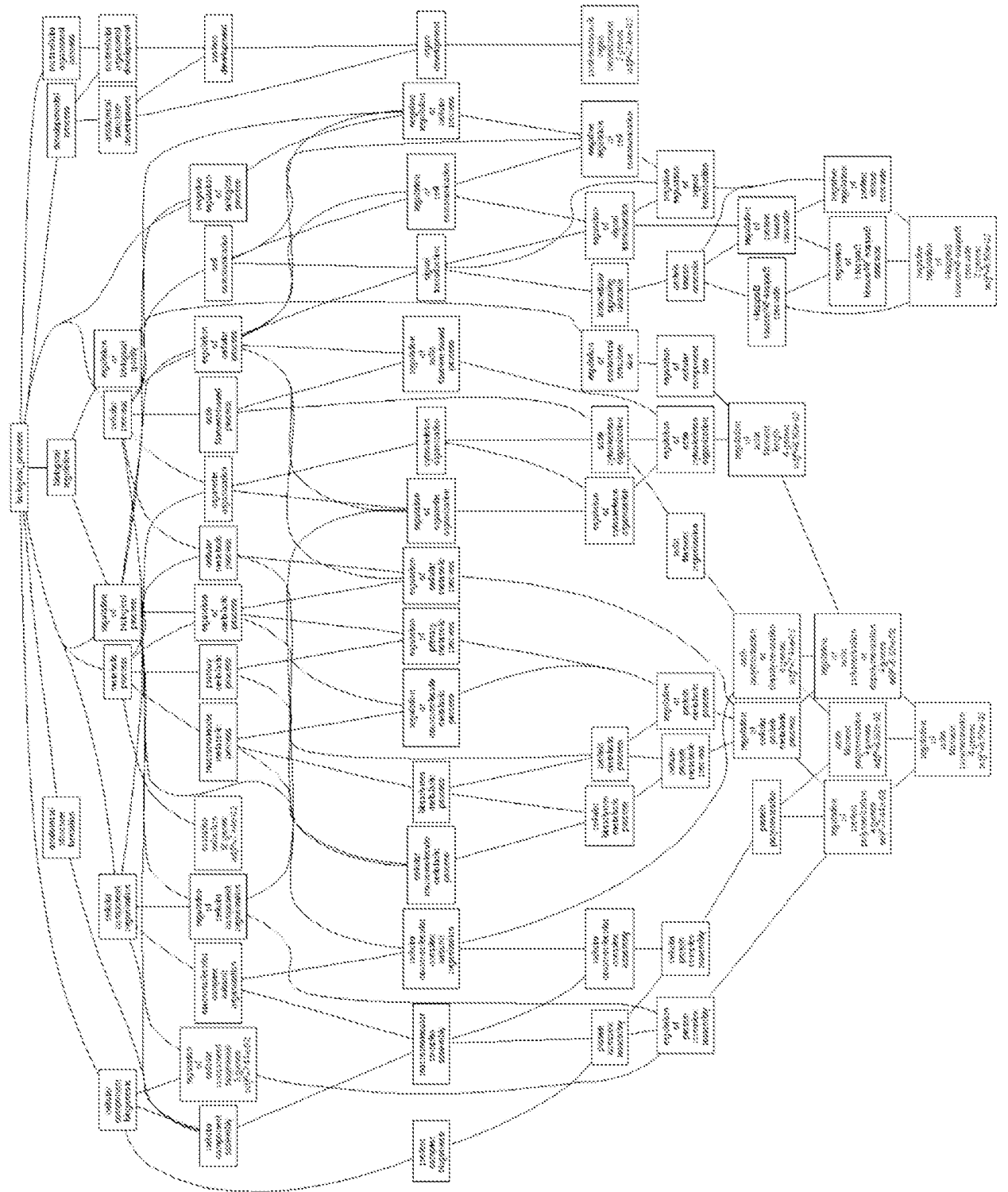


Figure S3

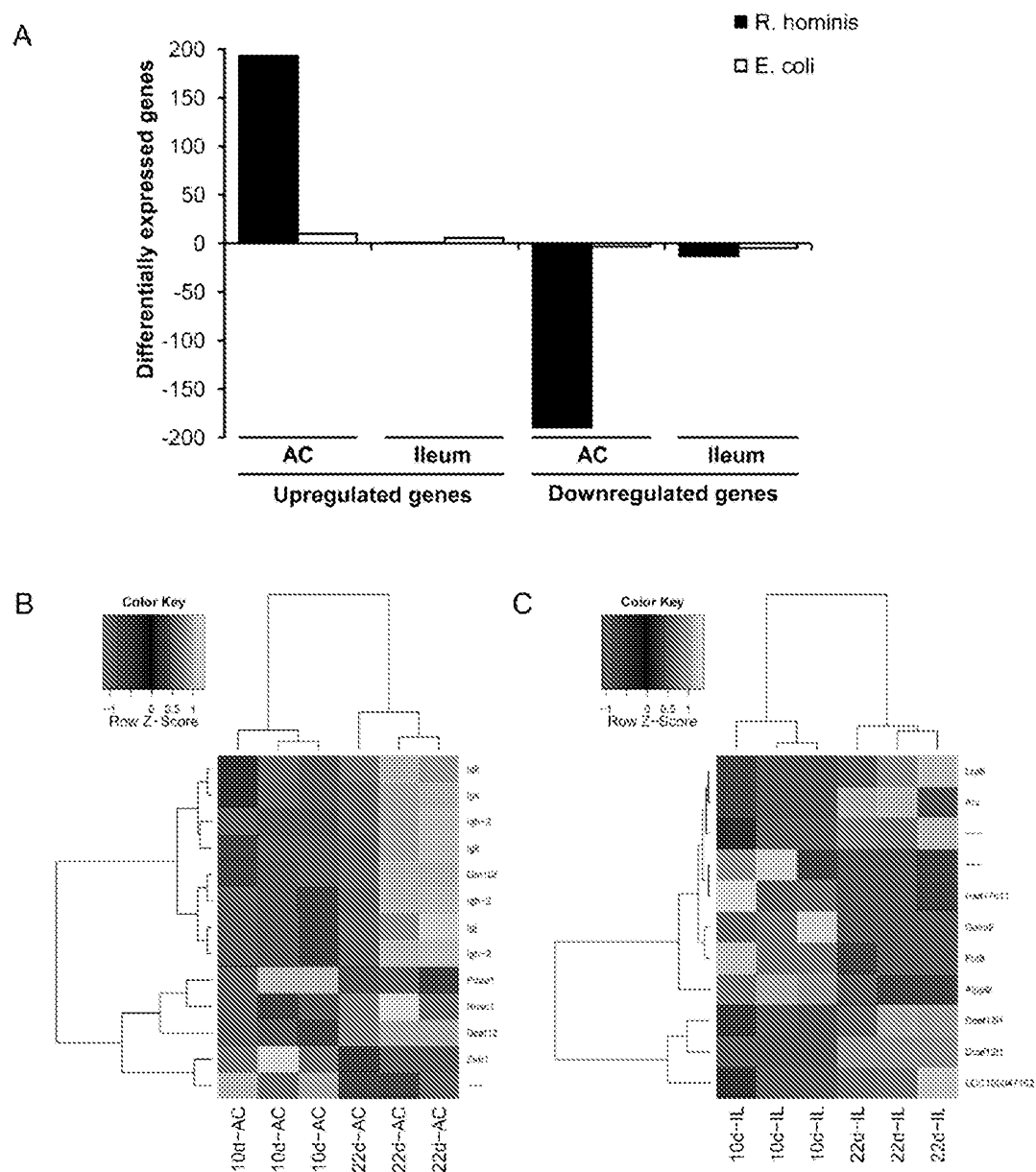


Figure S4

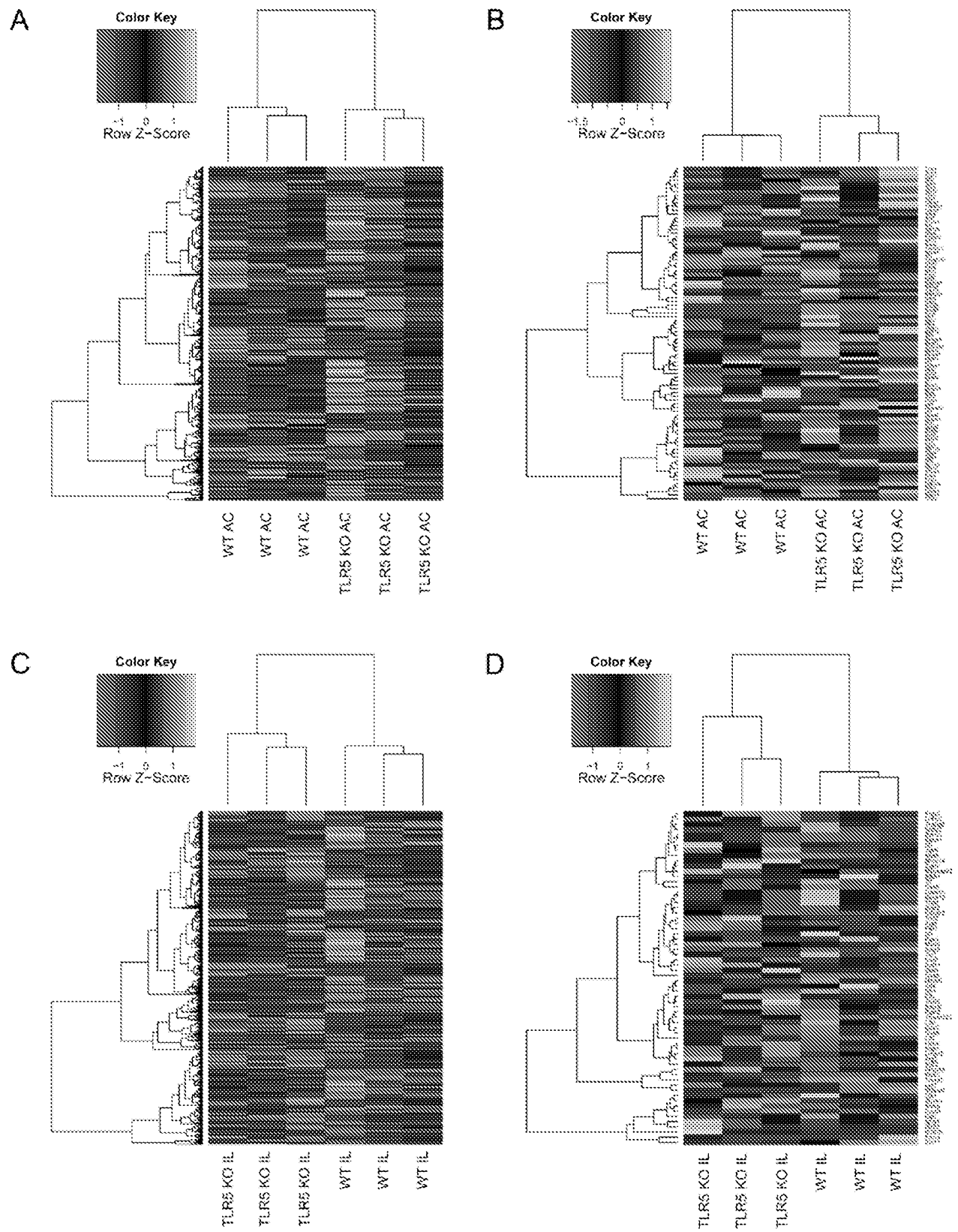


Figure S5

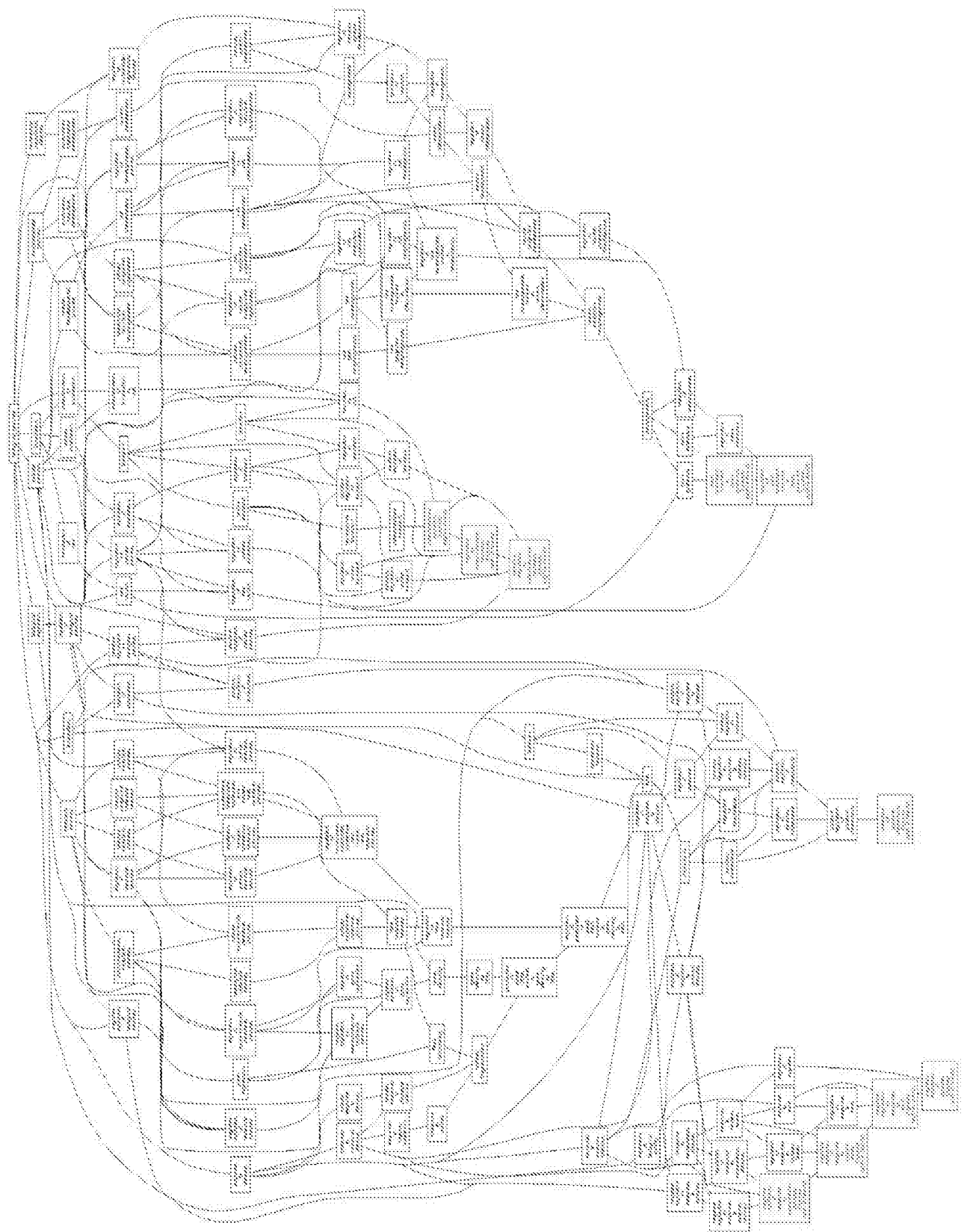


Figure S6

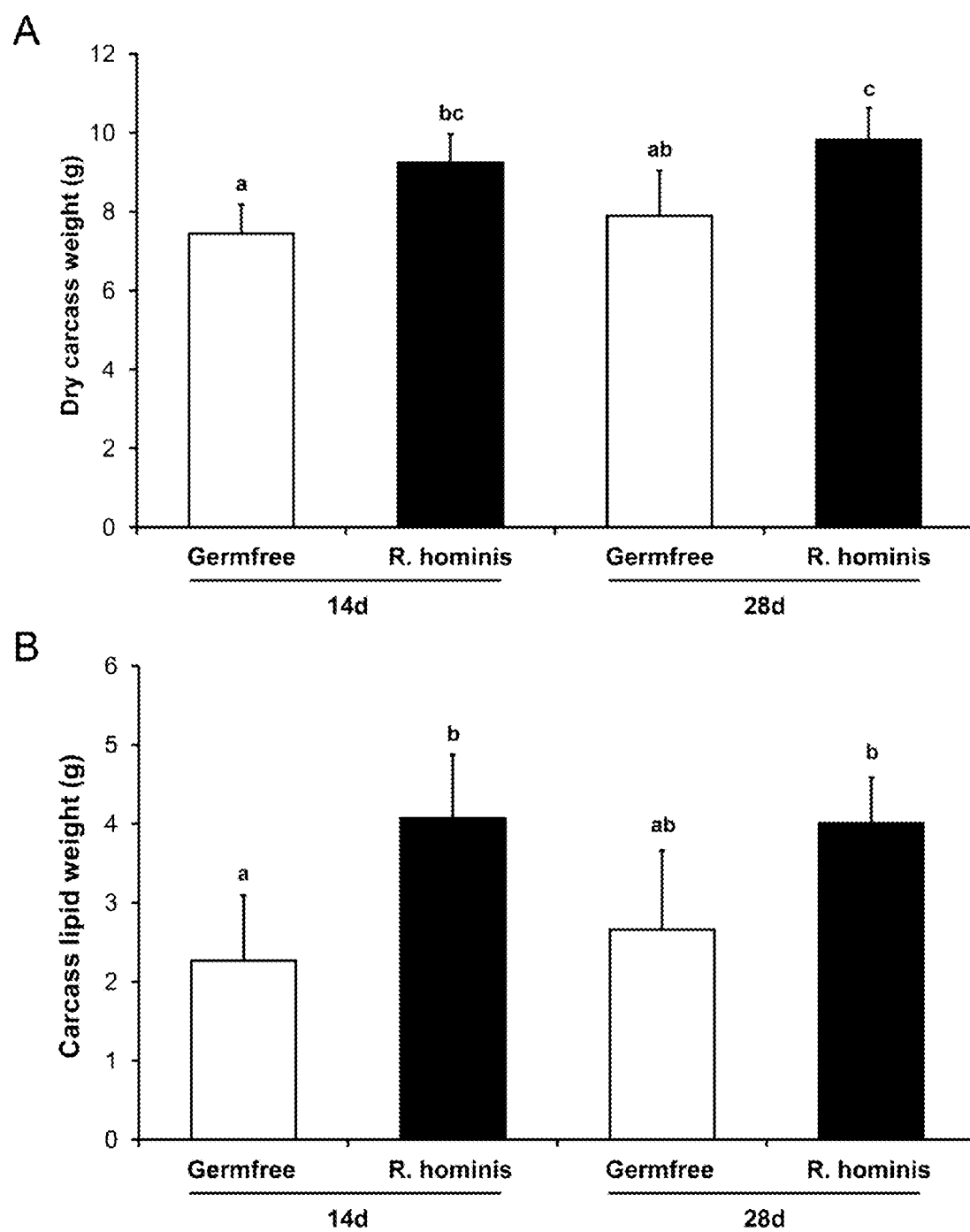


Diagram of the pCR*-Blunt II-TOPO* 3.5 kb circular plasmid. The plasmid contains a pUC origin, a tetracycline resistance gene (TetR), and a TOPO* origin. Restriction sites for various enzymes are indicated around the plasmid: Not I, Hind III, Age I, Xba I, Eco RI, Sal I, Bam HI, Sma I, Eco K, Eco V, Eco T, Xba I, Not I, Xba I, Eco RI, and Age I. The plasmid is labeled "pCR*-Blunt II-TOPO*" and "3.5 kb".

The diagram illustrates the pT7-MAT-Tag-FLAG-2 plasmid construct. The linear DNA sequence at the top shows the following elements from left to right: RBS, ATG, MAT, MCS (Multiple Cloning Site), FLAG, and STOP. The MCS region is flanked by restriction enzyme sites: Hind III, Xba I, EcoR I, Sma I, Kpn I, and Not I. The circular plasmid map below shows the same elements in a circle. The pT7/lac promoter is located upstream of the MAT gene. The lacI gene is also present. The amp^r gene (ampicillin resistance) is located downstream of the FLAG gene. The pBR322 origin and f1 origin are also indicated. The plasmid is labeled pT7-MAT-Tag-FLAG-2 and has a size of 4838 bp.

Figure 3A

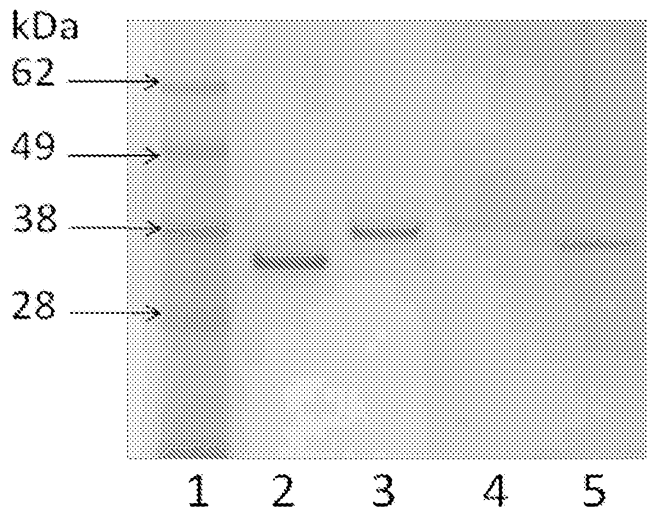


Figure 4A

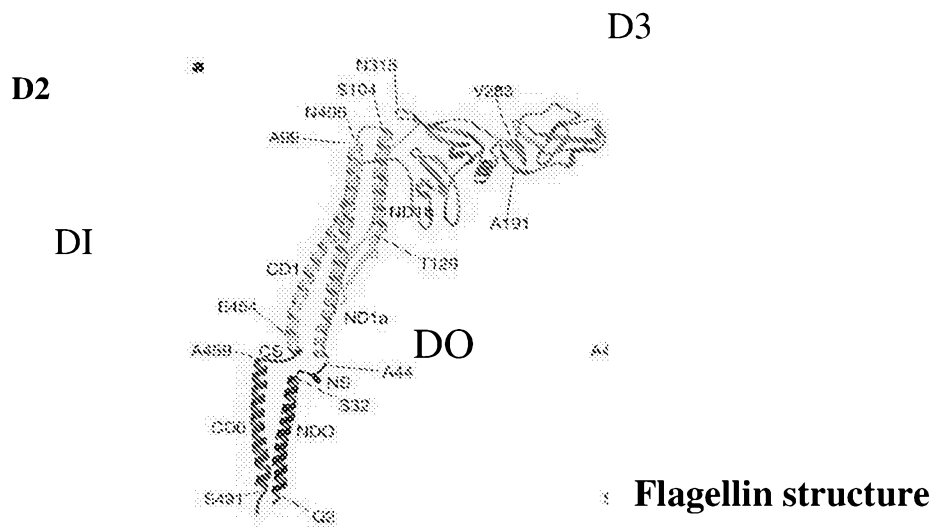


Figure B1 A&B

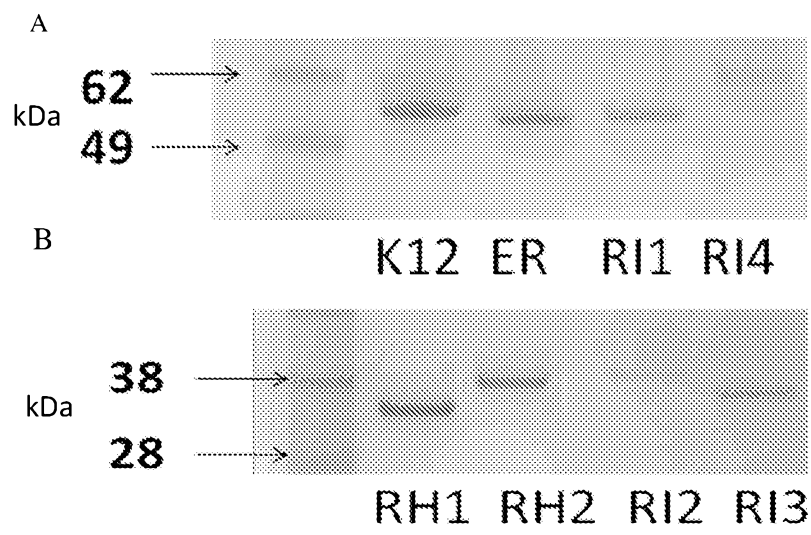


Figure C1

CLUSTAL 2.1 multiple sequence alignment confirming flagellin nucleotide sequence and accession number.

RIFla1	ATGCGTGGCGGAGACAATAGAAGGAGAAACAGAATGAGAATTAATTACAATGTGTCAGCA 60
RIEEV02820	ATGCGTGGCGGAGACAATAGAAGGAGAAACAGAATGAGAATTAATTACAATGTGTCAGCA 60
RIFla1	GCGATTGCGAATAAACATTTACTTGGAATTGAGGATAATTTAAGTGCATCGATGGAACGG 120
RIEEV02820	GCGATTGCGAATAAACATTTACTTGGAATTGAGGATAATTTAAGTGCATCGATGGAACGG 120
RIFla1	CTTTCATCGGGACTTAAGATCAACCATTCCAAGGACAATCCGGCAGGAATGGCTATTTCC 180
RIEEV02820	CTTTCATCGGGACTTAAGATCAACCATTCCAAGGACAATCCGGCAGGAATGGCTATTTCC 180
RIFla1	AACAAGATGAAAGCACAGATTGATGGTTTAAACCGGGCTTCCAGAATGCATCGGATGGT 240
RIEEV02820	AACAAGATGAAAGCACAGATTGATGGTTTAAACCGGGCTTCCAGAATGCATCGGATGGT 240
RIFla1	ATTTCTGTTATTCAGATCGCAGATGGTGCGCTGAGTGAAACGACCAGTATTTTACAGCGT 300
RIEEV02820	ATTTCTGTTATTCAGATCGCAGATGGTGCGCTGAGTGAAACGACCAGTATTTTACAGCGT 300
RIFla1	ATGAGAGAACTTTCCGTGCAGGCAGCGAGTGATGCAACAATGACACCGGCGGATAAAGAA 360
RIEEV02820	ATGAGAGAACTTTCCGTGCAGGCAGCGAGTGATGCAACAATGACACCGGCGGATAAAGAA 360
RIFla1	GCAATCCAGAAAGAAATCACTTCATTAAGATGAAGTTGACCGTATTTCTACAGATACA 420
RIEEV02820	GCAATCCAGAAAGAAATCACTTCATTAAGATGAAGTTGACCGTATTTCTACAGATACA 420
RIFla1	GAGTATAACAGCAAAACACTTTTAGATGGTTCATTAGATACCAGGGTTTACACCAAAAAT 480
RIEEV02820	GAGTATAACAGCAAAACACTTTTAGATGGTTCATTAGATACCAGGGTTTACACCAAAAAT 480
RIFla1	GCAACAAGAGTGGACATTTCTGATCATGTGAAAGCAGGACAGTATCAGCTTTCATTGAT 540
RIEEV02820	GCAACAAGAGTGGACATTTCTGATCATGTGAAAGCAGGACAGTATCAGCTTTCATTGAT 540
RIFla1	ACTGCAGCTACACAGGCCGGACCGGTAACAGCAAATCAGAATTATAATTCCACAGCACCG 600
RIEEV02820	ACTGCAGCTACACAGGCCGGACCGGTAACAGCAAATCAGAATTATAATTCCACAGCACCG 600
RIFla1	GTCGGTGCGTCCGGAACAATGAGTATTAATGGTTCTAAAGTAGAGATAGAGGCAGCCGAC 660
RIEEV02820	GTCGGTGCGTCCGGAACAATGAGTATTAATGGTTCTAAAGTAGAGATAGAGGCAGCCGAC 660
RIFla1	ACCTATGCGGAGGCTTTTGAGAAGATCAGAAATGCAGCAGAGACTGGTGAAACAACCGTT 720
RIEEV02820	ACCTATGCGGAGGCTTTTGAGAAGATCAGAAATGCAGCAGAGACTGGTGAAACAACCGTT 720
RIFla1	AAGATTGAAAAGAATGGAGCACTTTCATTTACCGCAGAACAGTACGGAATGTCAAGCATC 780
RIEEV02820	AAGATTGAAAAGAATGGAGCACTTTCATTTACCGCAGAACAGTACGGAATGTCAAGCATC 780
RIFla1	TTAGAGATCGCATTTGATGATAAGCAGCTTGCTAATGCATTGGATTTACAGCAGACGGA 840
RIEEV02820	TTAGAGATCGCATTTGATGATAAGCAGCTTGCTAATGCATTGGATTTACAGCAGACGGA 840
RIFla1	GGAAACAGTGTTGTAGAAGATCCAGAGAATAAAGGCAGCTATGTATACGGACAGATTGAG 900
RIEEV02820	GGAAACAGTGTTGTAGAAGATCCAGAGAATAAAGGCAGCTATGTATACGGACAGATTGAG 900
RIFla1	AATGGCAAAGTGATCGTACCTTCCGGTACAGATGCCGAAGTAACGCTCACAAAACCGAGT 960
RIEEV02820	AATGGCAAAGTGATCGTACCTTCCGGTACAGATGCCGAAGTAACGCTCACAAAACCGAGT 960
RIFla1	GATGGAACCGGATTTGGTGATACAGCTACGGTAAAAACAGATGGAAATAAGATTACGGTT 1020
RIEEV02820	GATGGAACCGGATTTGGTGATACAGCTACGGTAAAAACAGATGGAAATAAGATTACGGTT 1020
RIFla1	ACAGACAGAGCCGGATTTGAGATGTCATTTCTTGCTGATGCAGGTTATACGGGTAAGCTG 1080
RIEEV02820	ACAGACAGAGCCGGATTTGAGATGTCATTTCTTGCTGATGCAGGTTATACGGGTAAGCTG 1080
RIFla1	GATTTTGATGTCACGGATATCGGAACGATGGCACTTCATATTGGAGCAAATGAGGATCAG 1140
RIEEV02820	GATTTTGATGTCACGGATATCGGAACGATGGCACTTCATATTGGAGCAAATGAGGATCAG 1140
RIFla1	GAAACAAGAGTGCGTATTCCGGAGGTTTCTGCAAGAGCCTTTACATTGATGATGCAGAC 1200
RIEEV02820	GAAACAAGAGTGCGTATTCCGGAGGTTTCTGCAAGAGCCTTTACATTGATGATGCAGAC 1200

Figure C1 cont

RIFla1	GTGACGACTGTAAATGGAGCAGGCAGAGGTATCACACAGTTTGACGATGCCATTTCAAAG	1260
RIEEV02820	GTGACGACTGTAAATGGAGCAGGCAGAGGTATCACACAGTTTGACGATGCCATTTCAAAG	1260
RIFla1	GTCAGTGAAGTGC GTTCAAGACTTGGTGCATACCAGAATCGTCTTGAGAGTACGGTATCA	1320
RIEEV02820	GTCAGTGAAGTGC GTTCAAGACTTGGTGCATACCAGAATCGTCTTGAGAGTACGGTATCA	1320
RIFla1	AGCCTGGATACGTTTGAAGAAAATATGACAGGAGCCCAGTCACGACTGACAGATGCGGAT	1380
RIEEV02820	AGCCTGGATACGTTTGAAGAAAATATGACAGGAGCCCAGTCACGACTGACAGATGCGGAT	1380
RIFla1	ATGGCATCGGAAATGACAGATTATACACATCAGAATGTATTAATCAGGCAGCAATCTCT	1440
RIEEV02820	ATGGCATCGGAAATGACAGATTATACACATCAGAATGTATTAATCAGGCAGCAATCTCT	1440
RIFla1	GTTTTGACACAGGCAAACGATCTGCCACAGCAGGTATTGCAGATTCTGCAGTAA	1494
RIEEV02820	GTTTTGACACAGGCAAACGATCTGCCACAGCAGGTATTGCAGATTCTGCAGTAA	1494

Figure C2

CLUSTAL 2.1 multiple sequence alignment

RIFla2	ATGGTAGTTAATCATAATATGGCATTGATCTGTGAGAGTAGACAGTTACGATGTAATGTG	60
RIEEV02466	ATGGTAGTTAATCATAATATGGCATTGATCTGTGAGAGTAGACAGTTACGATGTAATGTG	60
RIFla2	AAGAACATGGAGAAGTCTTCAAAAAAGCTGGCAACAGGTTATAAATTGCTTGGAGCAAAT	120
RIEEV02466	AAGAACATGGAGAAGTCTTCAAAAAAGCTGGCAACAGGTTATAAATTGCTTGGAGCAAAT	120
RIFla2	GATGATGCAGCAGGATTACAGATATCAGAAACCATGCGTCATCAGACCAGAGGTCTTAAC	180
RIEEV02466	GATGATGCAGCAGGATTACAGATATCAGAAACCATGCGTCATCAGACCAGAGGTCTTAAC	180
RIFla2	AAAGCATCCAGAAATTCGCAAGATGGAATTAGTATGCTGCAGACAGCAGATGCAGCATT	240
RIEEV02466	AAAGCATCCAGAAATTCGCAAGATGGAATTAGTATGCTGCAGACAGCAGATGCAGCATT	240
RIFla2	CAGGAGACACAGGAAGTGTGGATCGAATGACGGATCTGACAACACAGGCAGCTAATGAT	300
RIEEV02466	CAGGAGACACAGGAAGTGTGGATCGAATGACGGATCTGACAACACAGGCAGCTAATGAT	300
RIFla2	ATCAATACGGATGCGGATCGTCGTGCAATTCAGGATGAAATCGATCAGTTAAATCAGGAA	360
RIEEV02466	ATCAATACGGATGCGGATCGTCGTGCAATTCAGGATGAAATCGATCAGTTAAATCAGGAA	360
RIFla2	GTGGATCGTATTGCATATACGACGAATTTAATCAGCAGTATATATTAGCGGATGGAAC	420
RIEEV02466	GTGGATCGTATTGCATATACGACGAATTTAATCAGCAGTATATATTAGCGGATGGAAC	420
RIFla2	CCGCAAGCAAGACCAGGATACTATATGATACAGACAGGAAGTCTTGCGGGACAGGGAATA	480
RIEEV02466	CCGCAAGCAAGACCAGGATACTATATGATACAGACAGGAAGTCTTGCGGGACAGGGAATA	480
RIFla2	GAGATTAAGTTTGTTAATGCGAGCAAAGAGAGCTTGGGTGTGGACAAGGTTGATGTATCA	540
RIEEV02466	GAGATTAAGTTTGTTAATGCGAGCAAAGAGAGCTTGGGTGTGGACAAGGTTGATGTATCA	540
RIFla2	TCGCATGCAAAAGCGACAGAATCTATAGCAGTGGTACAGAATGCAATTGAAAAGGCAGCT	600
RIEEV02466	TCGCATGCAAAAGCGACAGAATCTATAGCAGTGGTACAGAATGCAATTGAAAAGGCAGCT	600
RIFla2	TCGTTTAGAGATACATTTGGGGCACAACAGGAGCGGTTAGAACACGCATTGCGTGGAACG	660
RIEEV02466	TCGTTTAGAGATACATTTGGGGCACAACAGGAGCGGTTAGAACACGCATTGCGTGGAACG	660
RIFla2	GATAATACATCAGAAAGTACACAGAGGGCAGAATCAAGTAGACGCGATACCAACATGAAT	720
RIEEV02466	GATAATACATCAGAAAGTACACAGAGGGCAGAATCAAGTAGACGCGATACCAACATGAAT	720
RIFla2	ATGGAGATGGTACAATATTCTACAAACCGTATTTTAGTACAGGCATCTCAGAGTATTTTA	780
RIEEV02466	ATGGAGATGGTACAATATTCTACAAACCGTATTTTAGTACAGGCATCTCAGAGTATTTTA	780
RIFla2	GCACAGTACAATGATGATGCAAAATATGTGTTGGAAATGTTAAATAG	828
RIEEV02466	GCACAGTACAATGATGATGCAAAATATGTGTTGGAAATGTTAAATAG	828

Figure C3

CLUSTAL 2.1 multiple sequence alignment

RIFla3	ATGGTAGTACAGCACAAATATGACCGCAATGAATGCGAACAGAATGTTAGGCGTTACAACA 60
RIEEV00779	ATGGTAGTACAGCACAAATATGACCGCAATGAATGCGAACAGAATGTTAGGCGTTACAACA 60
RIFla3	AGCGCACAGGCAAAATCTTCAGAGAAATTATCTTCTGGTTACAGAATCAACCGTGCAGGT 120
RIEEV00779	AGCGCACAGGCAAAATCTTCAGAGAAATTATCTTCTGGTTACAGAATCAACCGTGCAGGT 120
RIFla3	GATGACGCTGCTGGTTTAAACAATTTCTGAGAAGATGAGAAGCCAGATCCGTGGATTAAAC 180
RIEEV00779	GATGACGCTGCTGGTTTAAACAATTTCTGAGAAGATGAGAAGCCAGATCCGTGGATTAAAC 180
RIFla3	AAAGCTTCTGACAACGCACAGGATGGTATTTCTTAATCCAGGTTGCTGAGGGTGCATTA 240
RIEEV00779	AAAGCTTCTGACAACGCACAGGATGGTATTTCTTAATCCAGGTTGCTGAGGGTGCATTA 240
RIFla3	TCTGAGACACATTCTATCTTACAGCGTATGAATGAGTTAGCTACTCAGGCTGCTAACGAT 300
RIEEV00779	TCTGAGACACATTCTATCTTACAGCGTATGAATGAGTTAGCTACTCAGGCTGCTAACGAT 300
RIFla3	ACCAATACAACCTGCTGATAGAGGAGCTATTCAGGATGAGATCAACCAGTTAACATCTGAG 360
RIEEV00779	ACCAATACAACCTGCTGATAGAGGAGCTATTCAGGATGAGATCAACCAGTTAACATCTGAG 360
RIFla3	ATTAACAGAATCTCTTCTACAACCTCAGTTCAATACTCAGAACCTCATCGATGGTACATTC 420
RIEEV00779	ATTAACAGAATCTCTTCTACAACCTCAGTTCAATACTCAGAACCTCATCGATGGTACATTC 420
RIFla3	GCAAATAAAAACCTTCAGGTTGGTTCTATCTGTGGACAGAGAATTACTGTTTCTATCGAC 480
RIEEV00779	GCAAATAAAAACCTTCAGGTTGGTTCTATCTGTGGACAGAGAATTACTGTTTCTATCGAC 480
RIFla3	AGTATGTCTGCTGGTAGCTTAAATGTATCTGCTAACTTAGTAAAGGTTAACACTTTCAGT 540
RIEEV00779	AGTATGTCTGCTGGTAGCTTAAATGTATCTGCTAACTTAGTAAAGGTTAACACTTTCAGT 540
RIFla3	GCAGCAGGTGAAGCAATGTCCAATATTCAGGGTGCTATTTCTGCAATTTCTACACAGCGT 600
RIEEV00779	GCAGCAGGTGAAGCAATGTCCAATATTCAGGGTGCTATTTCTGCAATTTCTACACAGCGT 600
RIFla3	TCTTACTTAGGAGCTCTTCAGAATCGTCTGGAGCATAACAATCTCCAATTGGACAACATT 660
RIEEV00779	TCTTACTTAGGAGCTCTTCAGAATCGTCTGGAGCATAACAATCTCCAATTGGACAACATT 660
RIFla3	TCTGAGAATACTCAGTCTGCTGAATCTCGTATCCGTGATACAGATATGGCTGAAGAGATG 720
RIEEV00779	TCTGAGAATACTCAGTCTGCTGAATCTCGTATCCGTGATACAGATATGGCTGAAGAGATG 720
RIFla3	GTTACTTACAGCAAGAACAATATTCTTGCTCAGGCAGGACAGTCTATGCTTGCTCAGGCT 780
RIEEV00779	GTTACTTACAGCAAGAACAATATTCTTGCTCAGGCAGGACAGTCTATGCTTGCTCAGGCT 780
RIFla3	AACCAGTCTACTCAGGGTGACTTTCTCTGTTACAGTAA 819
RIEEV00779	AACCAGTCTACTCAGGGTGACTTTCTCTGTTACAGTAA 819

Figure C4
CLUSTAL 2.1 multiple sequence alignment

RIFla4	ATGGCAATGGTAGTACAGCACAAACATGTCCGCAATGAATGCGAACAGAAATTTAGGTGTT 60
RIEEU99488	ATGGCAATGGTAGTACAGCACAAACATGTCCGCAATGAATGCGAACAGAAATTTAGGTGTT 60
RIFla4	ACAACAGGAATGCAGGCAAAATCATCAGAGAAGTTATCTTCCGGTTACAAGATCAACCGT 120
RIEEU99488	ACAACAGGAATGCAGGCAAAATCATCAGAGAAGTTATCTTCCGGTTACAAGATCAACCGT 120
RIFla4	GCAGCAGATGATGCAGCAGGACTTTCTATTTCTGAGAAGATGAGAAGCCAGATCCGCGGT 180
RIEEU99488	GCAGCAGATGATGCAGCAGGACTTTCTATTTCTGAGAAGATGAGAAGCCAGATCCGCGGT 180
RIFla4	TTAAATAAAGCATCTGACAATGCACAGGATGGTATCTCTTTAATCCAGACCGCTGAGGGA 240
RIEEU99488	TTAAATAAAGCATCTGACAATGCACAGGATGGTATCTCTTTAATCCAGACCGCTGAGGGA 240
RIFla4	GCATTAAATGAGTCCCACTCTATTTTACAGAGAATGAGAGAGTTATCCGTACAGGCAGCC 300
RIEEU99488	GCATTAAATGAGTCCCACTCTATTTTACAGAGAATGAGAGAGTTATCCGTACAGGCAGCC 300
RIFla4	AACGGTACAGAGACAGATGACGACCGCGAGGCAGTACAGAACGAGGTTTCCAGTTACAG 360
RIEEU99488	AACGGTACAGAGACAGATGACGACCGCGAGGCAGTACAGAACGAGGTTTCCAGTTACAG 360
RIFla4	GAAGAGCTGACAAGAATTTCTGAGACAACAGAGTTCAACACGATGAAGCTGCTGGATGGT 420
RIEEU99488	GAAGAGCTGACAAGAATTTCTGAGACAACAGAGTTCAACACGATGAAGCTGCTGGATGGT 420
RIFla4	TCTCAGAGTGGAAGTACATCTTCAACCGGTCAGGTCCGAAGTTTGGTGTGTAGATGCA 480
RIEEU99488	TCTCAGAGTGGAAGTACATCTTCAACCGGTCAGGTCCGAAGTTTGGTGTGTAGATGCA 480
RIFla4	ACATTAGACGGTGCACTTGTAACATCTAACGTGAAAGGTATTAAAGTAGCAACAGCAGCT 540
RIEEU99488	ACATTAGACGGTGCACTTGTAACATCTAACGTGAAAGGTATTAAAGTAGCAACAGCAGCT 540
RIFla4	GCCACAACAACAAAGGCAGGTCAGGAGACTGCTATCTGGGCTGCTGATGGAAAGACATTA 600
RIEEU99488	GCCACAACAACAAAGGCAGGTCAGGAGACTGCTATCTGGGCTGCTGATGGAAAGACATTA 600
RIFla4	ACTTTAAATCTTTCGAAAAATAAGGTATATACACAGGACGAAATTGATGACTTGATCGCA 660
RIEEU99488	ACTTTAAATCTTTCGAAAAATAAGGTATATACACAGGACGAAATTGATGACTTGATCGCA 660
RIFla4	AATGCAAAACAGGAAGACAGTTCTGCAACGGGTGCACCGGCTGAAGTGAAAGTATCTTTA 720
RIEEU99488	AATGCAAAACAGGAAGACAGTTCTGCAACGGGTGCACCGGCTGAAGTGAAAGTATCTTTA 720
RIFla4	AAGAATGGTATTTTTAATGCAGATGCAGACACAACCTGCCGGAAGTGAACAGCCGGTGGT 780
RIEEU99488	AAGAATGGTATTTTTAATGCAGATGCAGACACAACCTGCCGGAAGTGAACAGCCGGTGGT 780
RIFla4	GTGAAGGCAGTATCTGATGAAGGAACAGTAACTGGATTTGTTGGTGCAGATACAATTTCA 840
RIEEU99488	GTGAAGGCAGTATCTGATGAAGGAACAGTAACTGGATTTGTTGGTGCAGATACAATTTCA 840
RIFla4	TTTACGGCAAATAAGTATGGAGCAGAGTTCAATGATACTGTATTTAAATTCAAATTTGAT 900
RIEEU99488	TTTACGGCAAATAAGTATGGAGCAGAGTTCAATGATACTGTATTTAAATTCAAATTTGAT 900
RIFla4	AAGGCAGCAGGCAAAGAAGAAGTAGAGACAAATACAGCAATTGAAATTGATGGAGCAAAT 960
RIEEU99488	AAGGCAGCAGGCAAAGAAGAAGTAGAGACAAATACAGCAATTGAAATTGATGGAGCAAAT 960
RIFla4	GCGGTAACAGCAGGTGAATATACAATTCATCTTGACGACGGCAAAGAATATACGGCAGAA 1020
RIEEU99488	GCGGTAACAGCAGGTGAATATACAATTCATCTTGACGACGGCAAAGAATATACGGCAGAA 1020
RIFla4	GATTTAGAAGATGTTCTTAAACCGGCAGGATTTCGACTTTGATGTTAAATTAAGTGGAAT 1080
RIEEU99488	GATTTAGAAGATGTTCTTAAACCGGCAGGATTTCGACTTTGATGTTAAATTAAGTGGAAT 1080
RIFla4	ACACCAGATGAGCCAAATACTTTATTTGCAACCAGTGGCGCATCAACTGTGACTGATATT 1140
RIEEU99488	ACACCAGATGAGCCAAATACTTTATTTGCAACCAGTGGCGCATCAACTGTGACTGATATT 1140
RIFla4	ACAATGGGTGCTGGCACC GCCGAGCTGGTCTTGGAAGTACAGATGCTATGTGGGGGCAA 1200
RIEEU99488	ACAATGGGTGCTGGCACC GCCGAGCTGGTCTTGGAAGTACAGATGCTATGTGGGGGCAA 1200

Figure C4 cont

RIFla4	GCTGGTTATGACAGT-TATCTTCTGGTGCTGGCATTACCTTGCAGATTGGTGCAAATGAA 1259
RIEEU99488	GCTGGTTATGACAGTGTATCTTCTGGTGCTGGCATTACCTTGCAGATTGGTGCAAATGAA 1260
RIFla4	GGTCAGACCATGAGTTTCTCTATCGATGACATGAGTGCAAGAGCACTTGGCGTAGATGGC 1319
RIEEU99488	GGTCAGACCATGAGTTTCTCTATCGATGACATGAGTGCAAGAGCACTTGGCGTAGATGGC 1320
RIFla4	AACAAAGTTGATTTAAGCACACAGGCTGGCGCACAGAAAGCAACTGATACCATTGATGCA 1379
RIEEU99488	AACAAAGTTGATTTAAGCACACAGGCTGGCGCACAGAAAGCAACTGATACCATTGATGCA 1380
RIFla4	GCAATCAAGAAAGTATCTGCACAGCGTGGTAGAATGGGTGCGATCCAGAACCGTCTGGAG 1439
RIEEU99488	GCAATCAAGAAAGTATCTGCACAGCGTGGTAGAATGGGTGCGATCCAGAACCGTCTGGAG 1440
RIFla4	CACACCATCAGCAACCTTGATACAGCAGCAGAGAATACCCAGACTGCAGAGTCCCGTATC 1499
RIEEU99488	CACACCATCAGCAACCTTGATACAGCAGCAGAGAATACCCAGACTGCAGAGTCCCGTATC 1500
RIFla4	CGTGATACAGATATGGCAGAAGAGATGGTTGAGTACTCCAAGAACAACATTCTTGACAG 1559
RIEEU99488	CGTGATACAGATATGGCAGAAGAGATGGTTGAGTACTCCAAGAACAACATTCTTGACAG 1560
RIFla4	GCAGGTCAGTCTATGCTTGACAGGCGAACCAGTCTACACAGGGTGTACTCTCCTTATTA 1619
RIEEU99488	GCAGGTCAGTCTATGCTTGACAGGCGAACCAGTCTACACAGGGTGTACTCTCCTTATTA 1620
RIFla4	CAGTAA 1625
RIEEU99488	CAGTAA 1626

Figure D.2

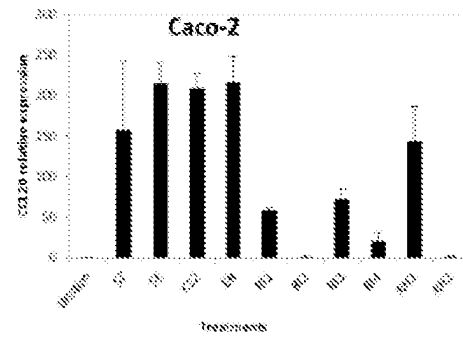
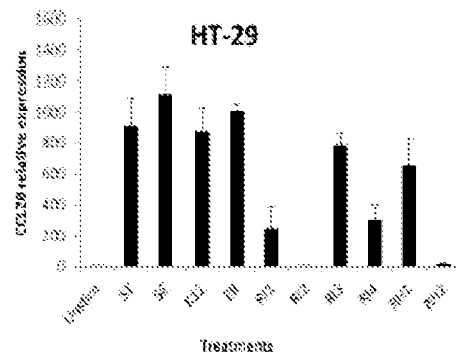


Figure D.3

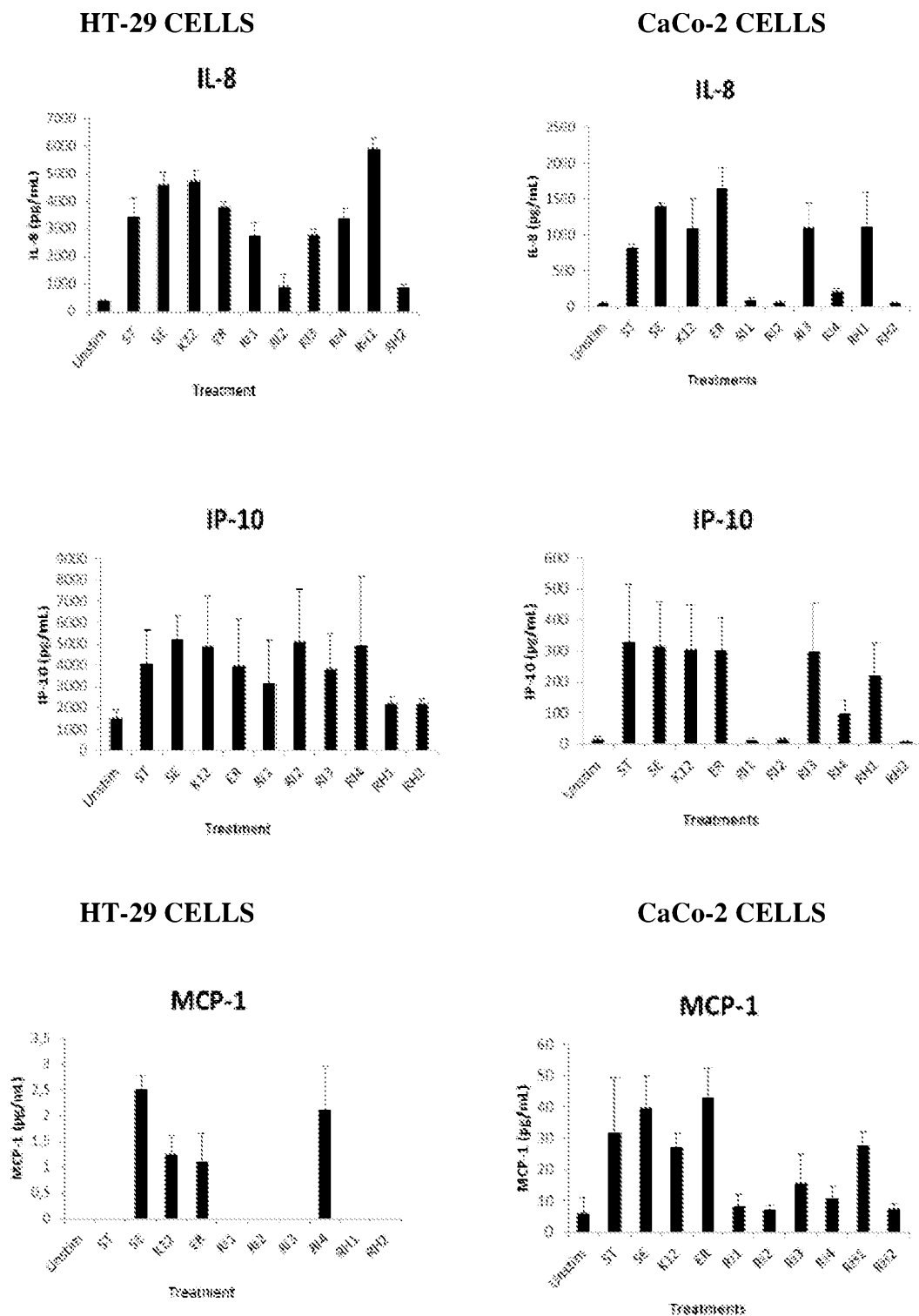
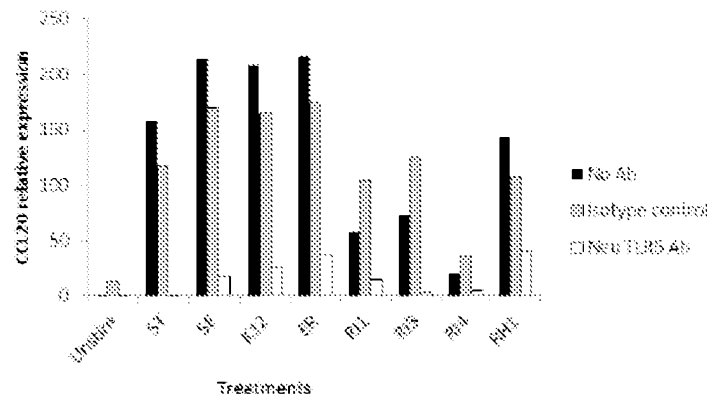


Figure D4



2014202021 10 Apr 2014

Figure D5A

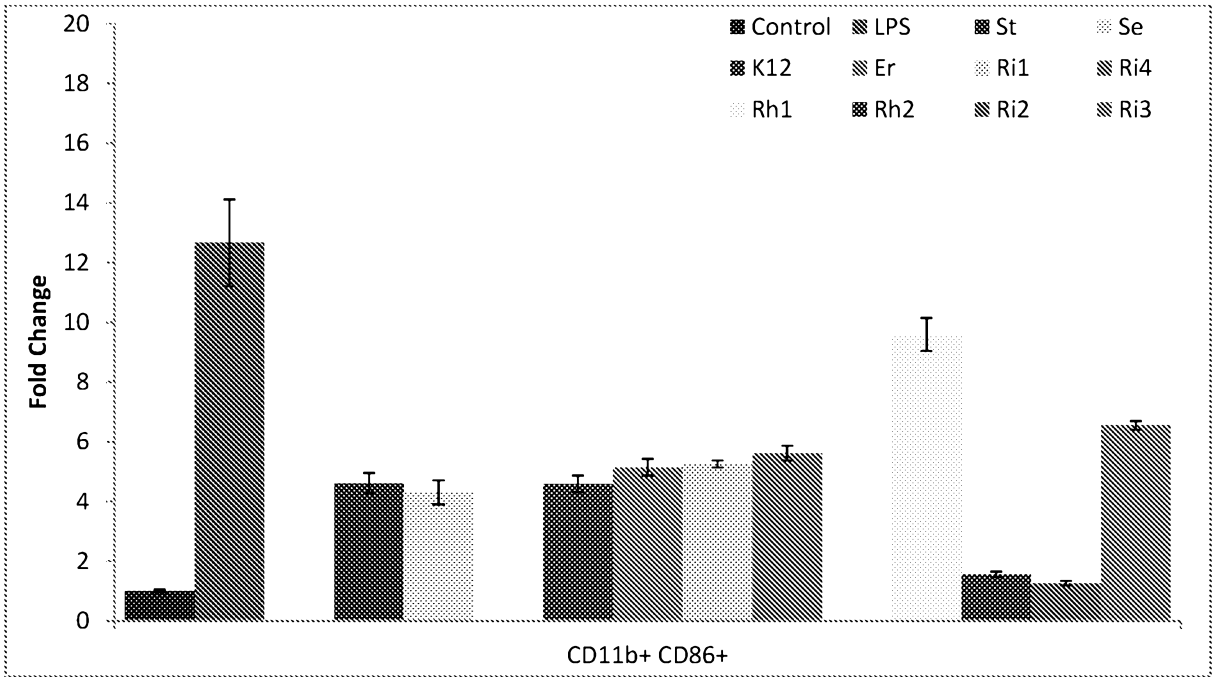


Figure D5B

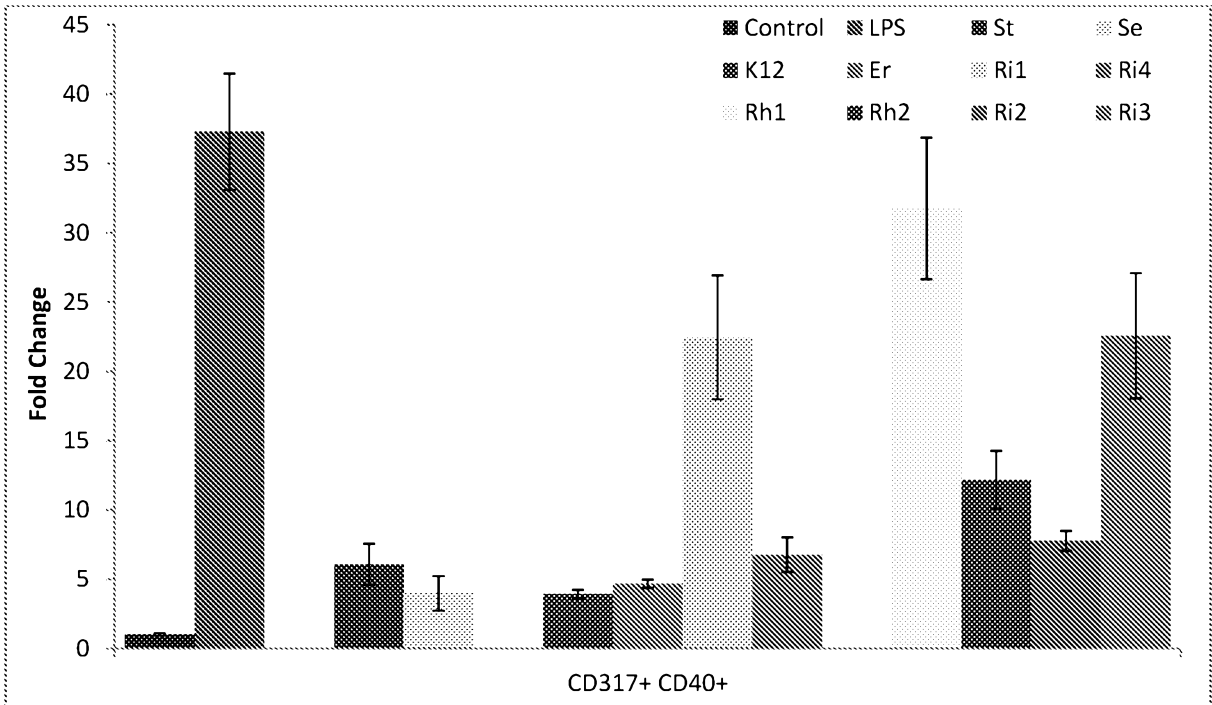
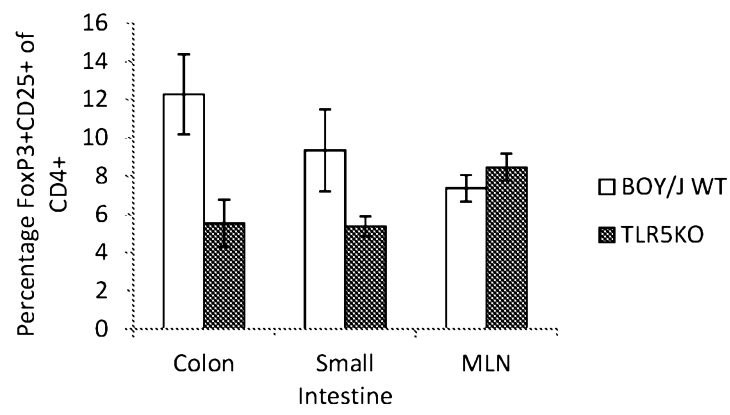


Figure D6

A



B

