Gut health: predictive biomarkers for preventive medicine and development of functional foods

Raish Oozeer1, Maria Rescigno2, R. Paul Ross3, Jan Knol4, Michael Blaut5, Artem Khlebnikov1 and Joël Dore6*

1Danone Research, RD 128, 91767 Palaiseau Cedex, France
2European Institute of Oncology, Milan, Italy
3Teagasc, Carlow, Republic of Ireland
4Danone Research, Wageningen, The Netherlands
5German Institute of Human Nutrition (DIFE), Nuthetal, Germany
6INRA, Domaine de Vilvert, 78352 Jouy-en-Josas, France

(Received 18 May 2009 – Revised 9 November 2009 – Accepted 13 November 2009 – First published online 7 April 2010)

There is an urgent need to develop and validate a series of biomarkers, which accurately measure and inform on how the human gut microbiota can affect human health. The human gut hosts a complex community of micro-organisms, with unique features in each individual. The functional role of this gut microbiota in health and disease is increasingly evident, but poorly understood. Comprehension of this ecosystem implies a significant challenge in the elucidation of interactions between all of its components, but promises a paradigm shift in preventive nutrition and medicine. ‘Oomics’ technologies for the first time offer tools of sufficient subtlety to tackle this challenge. However, these techniques must be allied with traditional skills of the microbial physiologist, which are in danger of being lost. Targeting these efforts at the identification of biomarkers associated with gut health will require access to a ‘biobank’ from a pan-European or worldwide observation study, which would include samples taken with appropriate frequency from healthy individuals of different ages. This offers a pragmatic opportunity for a unique food and pharmaceutical industry collaboration.

Intestinal microbiota: Predictive biomarkers: Gut health: Functional food

A workshop on the topic of Molecular biomarkers for the healthy gut microbiota was held in Bucharest, Romania, 4–6 February 2008. The workshop was convened by the Lactic Acid Bacteria Industrial Platform and was supported financially by the EU funded Specific Support Action GutImpact FP6-2003-517821. The aim of the present workshop was (1) to establish the state of the art of validated biomarkers and define expectations for ideal ones, (2) to give a prospective view of current developments of high throughput and molecular diagnostic methods and (3) to diminish the gap between molecular methods, biomarkers identification and strategies for product developments. The present report summarises the discussion that took place at the workshop.

What is the gut microbiota?

It is widely acknowledged that microbes residing in the human gut (the gut microbiota) are hugely significant in terms of human health. Various estimates suggest that the microbiota weighs about 1-5 kg and comprises approximately 95% of the total number of cells in the body. They are responsible for degrading certain food components and endogenous mucins, producing vitamins, preventing the colonisation or proliferation of pathogens and producing essential digestive and protective enzymes. As a result, they are thought to improve lactose intolerance, carry anti-tumour properties, neutralise certain toxins and reduce blood lipid levels. However, consequences of their metabolic potential may be the activation of substances with adverse health effects.

Gut microbes are both numerous and diverse. All three domains of life, Archaea, Bacteria and Eukarya, are represented. It is, however, the Bacteria that dominate the ecosystem, and in contrast to Archaea and Eukarya, they are present in everybody. Dominant bacteria that are regularly detected in intestinal contents of healthy individuals predominantly belong to four phyla: Actinobacteria, Bacteroidetes, Firmicutes and Verrucomicrobia. Many, very different types of bacteria occur within each of these, and over 420 species belonging to 117 different genera have been cultivated and identified to date. When it is appreciated that the huge majority (approximately 80%) of the gut microbiota cannot be cultured by conventional means, its
true complexity comes into focus\(^{7–9}\). Furthermore, in all those subjects analysed so far, it transpires that more than half of their gut microbes have not been found anywhere else and a large proportion represents totally novel taxonomic groups\(^{9–12}\). It seems therefore inevitable that further analysis of microbes from human gut samples will reveal novel phyla, classes, orders, families, genera and species of microbes.

The gut microbiota of each individual and its stability over time is thus a unique signature, as exquisite as the human fingerprint\(^{13,14}\). Extrapolation from available data suggests that thousands if not tens of thousands of different microbial species will eventually be assigned to it. In spite of the large number of bacterial species, it is also becoming clear that a significant proportion of dominant bacterial cells belong to a little more than sixty core species highly conserved between individuals\(^{15}\).

**Development and stability**

The intestine of a normal fetus is sterile. During birth and immediately afterwards, bacteria from the mother and the surrounding environment colonise the infant gut. Immediately after vaginal delivery, babies have bacterial strains in the upper gastrointestinal tract derived from the mothers’ colonic microbiota. Following birth, environmental, oral and cutaneous bacteria are readily transferred from the mother to the infant through suckling, kissing and caressing\(^{16}\). Within a few days, bacterial numbers reach \(10^9–10^{10}/g\) faeces\(^{17}\). During the first week of life, these bacteria create favourable conditions for the subsequent bacterial succession of strict anaerobes mainly belonging to the genera *Bifidobacterium*, *Bacteroides*, *Clostridium* and *Ruminococcus*. Breast-fed babies become dominated by bifidobacteria, possibly due to the presence of bifidobacterial growth factors in breast milk\(^{18}\). In contrast, the microbiota of formula-fed babies is more diverse with high numbers of *Enterobacteriaceae*, enterococci, bifidobacteria, *Bacteroides* and clostridia. After the introduction of solid food and weaning, the microbiota of formula-fed infants becomes similar to that of breast-fed infants. By the second year of life, the faecal microbiota resembles that of adults. Again, there is a characteristic shift in the composition of the intestinal microbiota upon ageing. In elderly persons, bifidobacteria decrease or become subdominant, while lactobacilli, enterococci, enterobacteria and clostridia increase\(^{19}\). This may be due to physiological and physico-chemical alterations of the digestive environment with ageing. This in turn may lead to increased risk of infection or degenerative diseases such as cancer or disorders of liver function.

**Relationship with human health**

Intestinal health is a state of well-being of the bowel and associated organs that is perceived at all stages of daily life and activities including defecation and is not merely the absence of intestinal discomfort. Bacteria in the gut fulfil a host of useful functions for human subjects, including digestion of energy-yielding substrates stimulating cell growth; repression of harmful micro-organisms; training the immune system to respond only to pathogens; defence/protection against some diseases. Hence, distortions in the composition of the bacterial community or their expressed functionalities can be associated with pathological conditions\(^{20–22}\).

**Digestion**

Without gut microbiota, the human body could not utilise some of the carbohydrates it consumes. Members of the microbiota have enzymes absent from human cells thus enabling them to break down certain non-digestible poly-saccharides (e.g. dietary fibre\(^{23}\)). They convert these into SCFA that can be used by host cells, providing a major source of useful energy and nutrients.

The SCFA are produced by a form of fermentation called saccharolytic fermentation and include mainly acetic, propionic and butyric acids\(^{24}\). These can have important physiological properties as follows: acetic acid is used by muscle; propionic acid helps the liver to produce ATP and may be used for gluconeogenesis; butyric acid provides energy to gut epithelial cells. In addition, the SCFA help the body to absorb nutrients such as Ca, Mg and Fe. Another benefit of SCFA is that they increase growth of intestinal epithelial cells and control their proliferation and differentiation. Bacterial cells also modulate intestinal growth by changing the expression of cell surface proteins such as Na-dependent glucose transporters (Na\(^+\).GLUT). In addition, interaction between bacteria and gut epithelial cells may prevent injury to the gut mucosa.

**Repression of pathogenic microbial growth**

Another important role of the microbiota is termed the ‘barrier effect’\(^{25}\). Growth of yeasts and harmful bacterial species such as *Clostridium difficile* is limited by competition from the host microbiota. The barrier effect protects human subjects from both invading species and species normally present in the gut at low numbers, whose growth is inhibited by the dominant gut microbiota.

The normal dominant microbiota contains bacterial strains that prevent the growth of pathogenic strains by competing for nutrients and attachment sites to the epithelium of the intestine. Symbiotic bacteria are favoured by this ecological niche and are thus more successful in the competition. The commensal bacteria may send chemical signals to the host about the type and amount of nutrients they may use, and the host responds accordingly\(^{25}\), thereby possibly starving out harmful bacteria. Commensal bacteria may also produce bacteriocins, substances harmful to other microbes, which may offer them an ecological advantage\(^{26}\).

**Immunity**

The microbiota has a dynamic effect on the host’s gut and systemic immune systems\(^{27,28}\). Bacteria are thought to be key in promoting early development of the gut’s mucosal immune system and continue to play a role in its operation later in life\(^{28}\). Bacteria stimulate the lymphoid tissue associated with the gut mucosa to produce antibodies to both commensals and pathogens\(^{29}\). The immune system recognises and counteracts harmful bacteria, but leaves the
helpful species alone, thanks to a tolerance mechanism most likely developed in infancy.

Antibodies received from the mother via the placenta or breastfeeding have the manifold role of imprinting the specific B-cell repertoire, protecting the newborn and shaping the microbiota (30). The first bacteria to colonise after birth are able to affect the immune response, making it more favourable to their own survival and less so to competing species; thus, the first bacteria to colonise the gut are important in determining the person’s lifelong gut microbiota make-up.

Recent findings have shown that gut bacteria play a role in the expression of toll-like receptors in the intestine, molecules that help the host to repair damage due to injury. Toll-like receptors cause parts of the immune system to repair injury caused, for example, such as that by irradiation.

Bacteria can influence the phenomenon known as oral tolerance, in which the immune system is less sensitive to an antigen (including those produced by gut bacteria) once it has been ingested. This tolerance, mediated in part by the gastrointestinal immune system and in part by the liver, can reduce an over-reactive immune response like those found in allergies and autoimmune diseases (31).

Some species of gut microbiota, such as some of those in the Bacteroides genus, are able to change their surface receptors thereby evading immune responses (32). Bacteria with neutral or harmful effects on the host may also have evolved these types of strategies. The host immune system has also adapted to this activity, preventing overgrowth of harmful species via the release of specific antibodies that down regulate the expression of epitopes involved in driving inflammatory responses (33,34).

**Preventing allergy**

Bacteria are implicated in preventing allergies, an over-reaction of the immune system to non-harmful antigens. Studies on the gut microbiota of infants and young children have shown that those who have allergies have different gut microbiota from those without allergies (35), with higher incidence of the harmful species C. difficile and Staphylococcus aureus and lower prevalence of Bacteroides and Bifidobacterium. One explanation is that since helpful gut microbiota stimulate the immune system and ‘train’ it to respond properly to antigens, a lack of these bacteria in early life leads to an inadequately trained immune system, which overreacts to harmless antigens (36).

**Preventing inflammatory bowel disease**

Another indication that commensal bacteria help to educate the immune system is the epidemiology of inflammatory bowel disease (IBD), such as Crohn’s disease. There is some evidence, for example, that SCFA reduce symptoms of IBD. In addition, some strains and species of bacteria can prevent inflammation. The incidence and prevalence of IBD are high in industrialised countries with a high standard of living and in less economically developed countries; incidence has increased in developed countries throughout the twentieth century. The disease is also linked to good hygiene in youth; lack of breastfeeding; and consumption of large amounts of sucrose and animal fat. Its incidence is inversely linked with poor sanitation during the first years of life and consumption of fruits, vegetables and unprocessed foods. In addition, the use of antibiotics, which disrupt homeostasis of the native gut microbiota along with harmful infectious pathogens, especially during childhood, is associated with IBD. On the other hand, using probiotics (bacteria consumed as part of the diet that imparts health benefits aside from just nutrition) can help to alleviate symptoms of certain forms of IBD (37).

**Factors affecting composition of the microbiota: the development of biomarkers**

The genotype of the host is thought to have the major influence on an individual’s microbiota. While the genome of an adult is fixed, the composition of its microbiota may alter with changes in lifestyle, diet and age. Molecular approaches used to examine the individuality and stability of the microbiota over time and to detect shifts in its composition after weaning could lead to the identification of biomarkers. In medicine, a biomarker is defined as an indicator of a particular disease state or a particular state of an organism. Specifically, a biomarker is a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention.

The scope for the identification of molecular biomarkers has been greatly enhanced by ‘omics’ technologies. These can unlock vast amounts of information stored in human genes (genomics); their products i.e. expressed RNA (transcriptomics) and translated proteins (proteomics); as well as the many small molecules participating in metabolic processes (metabolomics). These approaches are now being applied to understand the role of the gut microbiome in human disease using genomic analyses of microbial species from ‘humanised’ animal models and patient samples (38). There is some evidence, for example, that aberrant compositional development of the gut microbiota precedes weight gain/obesity (21), offering new possibilities for preventive and therapeutic applications in weight management (39).

Elsewhere, transcriptomic analysis has been used to understand how interactions among specific gut microbial species can affect host energy balance (40). For these analyses, a humanised mouse model was used, mice were raised so that the gut was free of microbes and then colonised with gut microbes commonly found in human subjects (41). Two microbial species in particular – Methanobrevibacter smithii and Bacteroides thetaiotaomicron – were found to have a cooperative relationship in digesting fibre as observed long ago for the rumen microbiota. That leads to more efficient intestinal nutrient absorption and energy storage as fat in a simplified mouse model. Further research building on the present work could lead to new ways to address both obesity and undernourishment.

To identify molecular biomarkers for a healthy gut microbiota, it is appropriate to further advance the delineation of criteria to define a normal homeostatic microbiota and further consider three sets of interactions: *microbe–food, microbe–host and microbe–microbe* interactions.
Microbe–food interactions

Fundamental questions underpinning the identification of biomarkers indicative of these interactions include: how does the microbiota respond to different dietary loads? What is the long-term influence of the impact of food on early colonisation? Is there imprinting from maternal feeding during pregnancy? What is the influence of a high-fat diet on the microbiota? How much energy is made available to the host via SCFA produced by bacteria? How does competition work between host and microbiota for various food constituents (sugars and amino acids) and how significant is it in terms of energy supply? Is it possible to identify key molecules of bacterial origin in blood and/or urine?

Several observations indicate that this is a fertile area for development. For example, it has been established that maternal milk contains components that are ‘active’ towards the microbiota. In addition, a high-fat diet is known to induce a change in the microbiota and an increased production of endotoxin. It is also possible to stimulate bifidobacteria with fructo-oligosaccharides.

Catabolic pathways have been elucidated for major dietary ingredients such as carbohydrates and proteins, and in some cases, contributing micro-organisms have been identified from in vitro studies of isolates. It is also clear that individual differences in metabolic pathways exist and again, indicator organisms can be identified. In either case, indicator metabolites might be appropriate alternatives. A metabolic model of food-compound degradation with predictive value using environmental metabolomics and metagenomics data should be the ultimate objective.

Microbe–host interactions

The focus of activity of microbiota–host interactions is the gut barrier or mucosa. The normal colonisation of the mammalian intestine with commensal microbes is thought to drive the development of the host immune systems during neonatal life and to maintain the physiologically steady state of immune stimulation in the gut throughout life. Therefore, biomarkers might be sought in mucus, epithelial cells and other immunocompetent cells of the gut-associated lymphoid tissues, and among anti-microbial peptides, secreted antibodies and receptors of microbe-derived signals.

Questions that need to be answered include: should the biomarkers be sought for locally in the gut or systemically? Can a systemic biomarker represent a locally disregulated response? Is it possible to identify systemic biomarkers that can predict the onset of a disease, or is it possible that when they reach the systemic level it is already too late and there is overt disease?

On the microbe side, a ‘reverse science’ approach should be adopted, using bioinformatics tools (existing or novel) in order to identify genes that might code for a microbe’s surface proteins. If a surface protein elicits an immune response of interest, it will represent a potential biomarker.

Microbe–microbe interactions

Any given microbiota will contain about 1000 different species of dominant bacteria. Given this complexity, it is undoubtedly that bacterial communication and antagonism has a dominant role in defining the microbial composition of the gut. Consequently, it is vital that biomarkers that can inform on or predict microbe–microbe interactions need to be identified. These biomarkers will be developed based on mechanisms by which bacteria communicate and affect each others performance. These could include:

1. Quorum-sensing where signals are produced by one bacterium which are interpreted by its neighbour.
2. Direct antagonism where a bacterium produces an anti-microbial metabolite that kills susceptible microbes.
3. Horizontal gene transfer mediated through conjugation, transduction and transformation.
4. Competitive exclusion whereby certain microbes predominate in particular niches.
5. Tolerance and resistance to anti-microbials produced by other gut microbes.
6. Competition for nutrients.
7. Oxidative species/redox balancing.
8. Bacterial lysis, which would provide nutrients for growth of opportunistic bacteria.

Initially, the most promising targets would be stable biomarkers such as microbial signatures, metabolites or nucleic acids. A metagenomic/metaproteomic profile is a pattern of the presence or absence of particular genes or proteins across a set of organisms. Such profiling of microbial community structure at all levels of integrated microbiomics would support the development of microbiome–microbe biomarkers, based on mechanisms of bacterial dominance, persistence and dynamics.

A biobank for the development of preventive medicine

The identification of suitable biomarkers for disease is hampered by the difficulty in defining precisely what the healthy status of an individual means. Biomarkers should first identify the health condition and assess how far an individual deviates from this condition. Meaningful investigations of any of these interactions and subsequent development of biomarkers for preventive medicine require access to large volumes of human sample material. The ideal solution would be the establishment of a global ‘biobank’.

While this represents a major undertaking, the potential returns make it attractive to both food and pharmaceutical industries in collaboration with governments. Such an initiative would analyse a population of healthy individuals at the beginning that might develop disease over the years of analysis. This would allow establishment and identification of the health status of a population and identification of biomarkers that are predictive of a particular disease. The first step for such an exercise is to define operational parameters. These include:

- Population cohorts, for example: ages 0–21; 60–81 months; specific baby groups (normal v. caesarean delivery; breastfed v. bottle fed).
- Population size. This should be large enough to include individuals that will develop a disease. For instance: if 9% of children develop asthma, the population number should be sufficiently high to reach statistical significance.
on the observed changed parameters in 9% of the population.

Nature of sampling procedures and material to be collected.

Standard protocols for sample collection would need to be agreed. It is important to promote the idea that minimal invasive techniques are acceptable and also that the associated ethical issues need to be addressed. Rectal biopsy is, for example, no more invasive than blood sampling at the physiological level. Sampling of individuals should be linked to data obtained from health/lifestyle questionnaires. Appropriate material for collection includes faeces, blood, saliva and urine.

**Nature of measurement.** Microbiological level (i.e. taxa from phylum to species and functional genes, proteins, metabolites). Cytological level (i.e. genes, RNA, proteins and metabolites).

**Timing of sampling.** At least one sampling per year, twice for babies aged 0–3 months.

Approaches such as high-throughput sequencing and tandem MS technologies are set to revolutionise measurement. The volumes of data generated will necessitate a parallel improvement in bioinformatics as more information will be generated than can be processed in classical ways. Improved tools need to be developed to integrate data and visualise it in metabolic maps, in order to build testable hypothesis.

Alongside these cutting edge developments, it is important not to neglect classical approaches. Expertise in bacterial physiology (especially for anaerobic metabolism) has been lost over time and tools must be developed to culture currently non-culturable micro-organisms.

**Attendees**

J. D., Institut national de la recherche agronomique, France; R. O. and A. K., Danone Research, France; Willem de Vos, Wageningen UR, The Netherlands; Fabrizio Arigoni, Nestlé R&D, Switzerland; J. K., Danone Research, The Netherlands; Fredrik Bäckdahl, Sahlgrenska University Hospital, Sweden; Akira Kusuhiro, Yakult Honsha European Research Center, Belgium; Sue McKay, Unilever, The Netherlands; Francisco Guarner, University Hospital Vall d’Hebron, Spain; Douwina Bosscher, Orafti, Belgium; Dan L. Dumitrascu, University of Medicine and Pharmacy, Romania; P. R., Teagasc, Ireland; M. B., Institute of Human Nutrition (DIFE), Germany; M. R., European Institute of Oncology, Italy. Science writer: Paul Wymer, UK.

**Acknowledgements**

We would like to thank Irene Lenoir for initiating the present project and Sadigatou Diallo for the logistics organisation. The workshop was supported financially by the EU funded Specific Support Action GutImpact FP6-2003-517821. Conflict of interest: R. O., J. K. and A. K. are employees of Danone Research, member of the Lactic Acid Bacteria Industrial Platform. P. R. is an employee of Moorepark Food Research Centre, Ireland. Other authors, M. R., M. B. and J. D., had no conflict of interest.

**References**