

A metagenomic insight into our gut's microbiome

Patricia Lepage,^{1,2} Marion C Leclerc,^{1,2} Marie Joossens,^{3,4,5} Stanislas Mondot,⁶ Hervé M Blottière,^{1,2} Jeroen Raes,^{3,4} Dusko Ehrlich,^{1,2} Joel Doré^{1,2}

¹INRA, MICALIS UMR1319, Jouy-en-Josas, France

²AgroParisTech, MICALIS UMR1319, Jouy-en-Josas, France

³Research Group of Bioinformatics and (Eco-) systems Biology, Department of Structural Biology, VIB, Brussels, Belgium

⁴Microbiology Unit, Department of Applied Biological Sciences, Vrije Universiteit Brussel, Brussels, Belgium

⁵IBD Research Group, TARGID, Department of Gastroenterology, KU Leuven, Leuven, Belgium

⁶CSIRO, Division of Livestock Industries, Queensland Biosciences Precinct, St Lucia, Queensland, Australia

Correspondence to

Dr Patricia Lepage, MICALIS, Building 405, Domaine de Vilvert, INRA, 78350

Jouy-en-Josas, France; patricia.lepage@jouy.inra.fr

Published Online First
23 April 2012

ABSTRACT

Advances in sequencing technology and the development of metagenomic and bioinformatics methods have opened up new ways to investigate the 10^{14} microorganisms inhabiting the human gut. The gene composition of human gut microbiome in a large and deeply sequenced cohort highlighted an overall non-redundant genome size 150 times larger than the human genome. The *in silico* predictions based on metagenomic sequencing are now actively followed, compared and challenged using additional 'omics' technologies. Interactions between the microbiota and its host are of key interest in several pathologies and applying meta-omics to describe the human gut microbiome will give a better understanding of this crucial crosstalk at mucosal interfaces. Adding to the growing appreciation of the importance of the microbiome is the discovery that numerous phages, that is, viruses of prokaryotes infecting bacteria (bacteriophages) or archaea with a high host specificity, inhabit the human gut and impact microbial activity. In addition, gene exchanges within the gut microbiota have proved to be more frequent than anticipated. Taken together, these innovative exploratory technologies are expected to unravel new information networks critical for gut homeostasis and human health. Among the challenges faced, the *in vivo* validation of these networks, together with their integration into the prediction and prognosis of disease, may require further working hypothesis and collaborative efforts.

INTRODUCTION

Shortly after the publication of the first human genome draft in 2001 it became obvious that a complete understanding of the human biology would only be fully assessed by combining the analysis of the host and its surrounding environment.¹ The human gastrointestinal tract hosts more than 100 trillion bacteria and archaea which together make up the gut microbiota. Archaea are prokaryotes and represent a third domain between bacteria and eukaryotic organisms. First described from extreme environments, archaea exist in a broad range of habitats and may contribute up to 20% of the Earth's biomass. Most of the cultivable species are members of two phyla, the Euryarchaeota and Crenarchaeota. In gut ecosystems, they are methanogenic, from the Euryarchaeota phylum. Even though the amount of bacteria in the human

gut outnumber human cells by a factor of 10^2 ,^{2,3} some finely tuned mechanisms allow these microorganisms to colonise and survive within the host in a commensalism relationship.⁴ This tolerance phenomenon is facilitated through physical separation of bacteria and host cells via the mucus layer, but also through modifications of antigenic moieties of the microbiota components to reduce their immunogenic properties and by direct modulation of localised host immune responses.^{5–7}

The human gut microbiota can be considered an organ within an organ⁸ that co-evolved with humans to achieve a symbiotic relationship leading to physiological homeostasis.⁹ The human host provides a nutrient-rich environment and the microbiota provides indispensable functions that humans cannot exert themselves, such as the production of some vitamins, digestion of complex polysaccharides¹⁰ and the shaping of the immunological environment.^{11,12} Commensal bacteria influence the normal development and function of the mucosal immune system, the induction of IgA and epithelial barrier tightening. Through the production of short-chain fatty acids, mainly acetate, propionate and butyrate, resident bacteria positively influence intestinal epithelial cell proliferation and differentiation, and also mediate various metabolic effects.¹³

Development of cultivation-independent methods based on 16S small subunit (SSU) rRNA gene sequence analysis rapidly expanded our knowledge about the diversity of the gastrointestinal tract microbiota. Only a decade after their first application to the human gut,¹⁴ the number of detected gastrointestinal tract phylotypes using molecular techniques has by far outnumbered the cultivated gut species. From more than 1200 microbes described, only 12% were recovered by application of both molecular and cultivation-based approaches, while the vast majority (~75%) were detected solely through SSU rRNA sequencing.¹⁵

Chaotic in the early stages of human life,^{16,17} the assembly of the human gut microbiota remains globally stable over time in healthy conditions in the absence of perturbation.^{18,19} The average total number of bacterial species was estimated to be close to 1000 per individual whereas 10 000 to 40 000 are predicted for the whole microbiota population.^{20,21} The restricted number of phyla in comparison to other ecosystems (four out of 50 existing phyla) has suggested a tight co-evolutionary history between the host and its

microbiota.^{22 23} This co-evolution is also observed at the species level and has led to the definition of a phylogenetic core composed of 66 highly prevalent and abundant species.²¹ Two per cent of one person's microbial species are shared by at least 50% of the population.

Remarkably, shifts in the bacterial makeup of the human gut microbiota have been associated with digestive tract dysfunctions such as inflammatory bowel disease (IBD), irritable bowel syndrome (IBS) and obesity.^{20 24–31} More than 10 years ago, the concept of dysbiosis or unbalanced composition of the intestinal microbiota, was introduced in the IBD research field.³² Although an impressive list of documented microbial alterations in patients with IBD has recently been reviewed,³³ the original question of whether dysbiosis is just a secondary phenomenon in IBD or truly causal remains.³⁴

Further insights into the human gut ecosystem are needed to comprehend the exact role of microbiota in health and disease. Because most of the bacteria inhabiting the gut are uncultivable, their functions cannot be inferred from composition data. Knowing which microbes are there is not sufficient. Essential questions are 'What is the genetic potential of the non-cultured bacterial fraction of the gut microbiota?' and 'What are these microbes really doing?'. Meta-omics aims to answer these questions. Metagenomics, which is based on whole microbiome genomic content, has started to put forward the microbial functionalities embedded into the human gut microbiota. Among the key outputs, this review addresses the total number of encoding genes per person or as a whole, enterotypes definition and global functional properties.

METAGENOMICS: TOWARDS A BETTER UNDERSTANDING OF THE HUMAN GUT MICROBIOME

Metagenomics was first described in 1998 by Handelsman and Rodon.^{35 36} It was defined as analysis of the collective genomes that are present in a defined environment or ecosystem, hence giving insight into functions of non-cultivated bacteria. It was first applied to the analysis of aquatic and soil ecosystems. The first golden standard for metagenomics was the use of metagenomic libraries. Large DNA insert metagenomic libraries were constructed using fosmids, cosmids or bacterial artificial chromosomes with single, low copy or copy control vectors. Total microbial DNA was extracted, separated in large DNA fragments using pulse-field gel electrophoresis before being cloned in host bacteria to generate a screening library.

The emergence of next generation sequencing technologies,³⁷ such as pyrosequencing, SOLiD ('Sequencing by oligonucleotide ligation and detection') or Illumina led to the development of sequencing projects covering a broader fraction of the microbial diversity present in the original sample. It then became possible to sequence the

inserts of clones of interest from metagenomic libraries rapidly and in a cost-effective manner but also to directly shotgun sequence the metagenomic DNA content of an ecosystem. The bioinformatics analysis of a sequence, so called annotation, from open reading frames finding to similarity searches in databases, enables researchers to identify major functions and genes that encode enzymes of clinical or industrial interest. The complete genome of particularly abundant microorganisms can also be reconstructed,^{38 39} providing a more comprehensive view of their biological potential (figure 1).

Nonetheless, the scale of these analyses needs to be adequate to permit the mapping of the complete human microbial metagenome, given the incredible abundance and complexity of the microbial populations found in the digestive tract. An average of 4.5 gigabases (Gb; ranging between 2 and 7.3 Gb) of sequence were generated for each faecal sample of the 1st MetaHIT study.⁴⁰ Considering an average size of bacterial chromosomes of 4.5 megabases (Mb; ranging from 0.6 Mb to over 10 Mb), the gene catalogue was therefore equivalent to that of some 1000 bacterial species, encoding about 3364 non-redundant genes for each person's microbiome, capturing most of the novelty.

The first sequence-based characterisation was obtained from samples of two American healthy volunteers⁴¹ and showed a significant enrichment of the gut microbiome in metabolic pathways related to metabolism of glycans, amino acids and xenobiotics; methanogenesis; and 2-methyl-D-erythritol 4-phosphate pathway-mediated biosynthesis of vitamins and isoprenoids. Shortly after, the metagenomic analysis of 13 healthy Japanese volunteers⁴² confirmed, at functional levels, the chaotic picture drawn from previous compositional studies of unweaned infants' gut microbiome.¹⁷ Nevertheless, a set of 136 gene categories was found to be prevalent in infant-type microbiome, clearly directed towards carbohydrate transport and metabolism facilitating nutrient uptake from milk. The adult-type microbiome was found to be commonly enriched with 237 gene families, suggesting host–microbiota co-evolution towards functionalities favouring energy harvest from diet (carbohydrate metabolism) and bacterial competition (antimicrobial peptide transporters). In addition, genes encoding the biosynthesis of flagella and related to chemotaxis were depleted, suggesting that commensal microbes lowered their immunogenic properties to persist into the gut.

Finally, among the first human gut metagenomic libraries were two large insert libraries generated from pools of faecal samples from six healthy people and six patients with Crohn's disease.^{29 43} Analysis of the corresponding gene content at the phylogenetic level indicated, as expected from previous works, broad species diversity in healthy people and, inversely in the Crohn's disease (CD) faecal library, a markedly reduced species diversity belonging to the Firmicutes, mostly from the *Clostridium leptum* group.

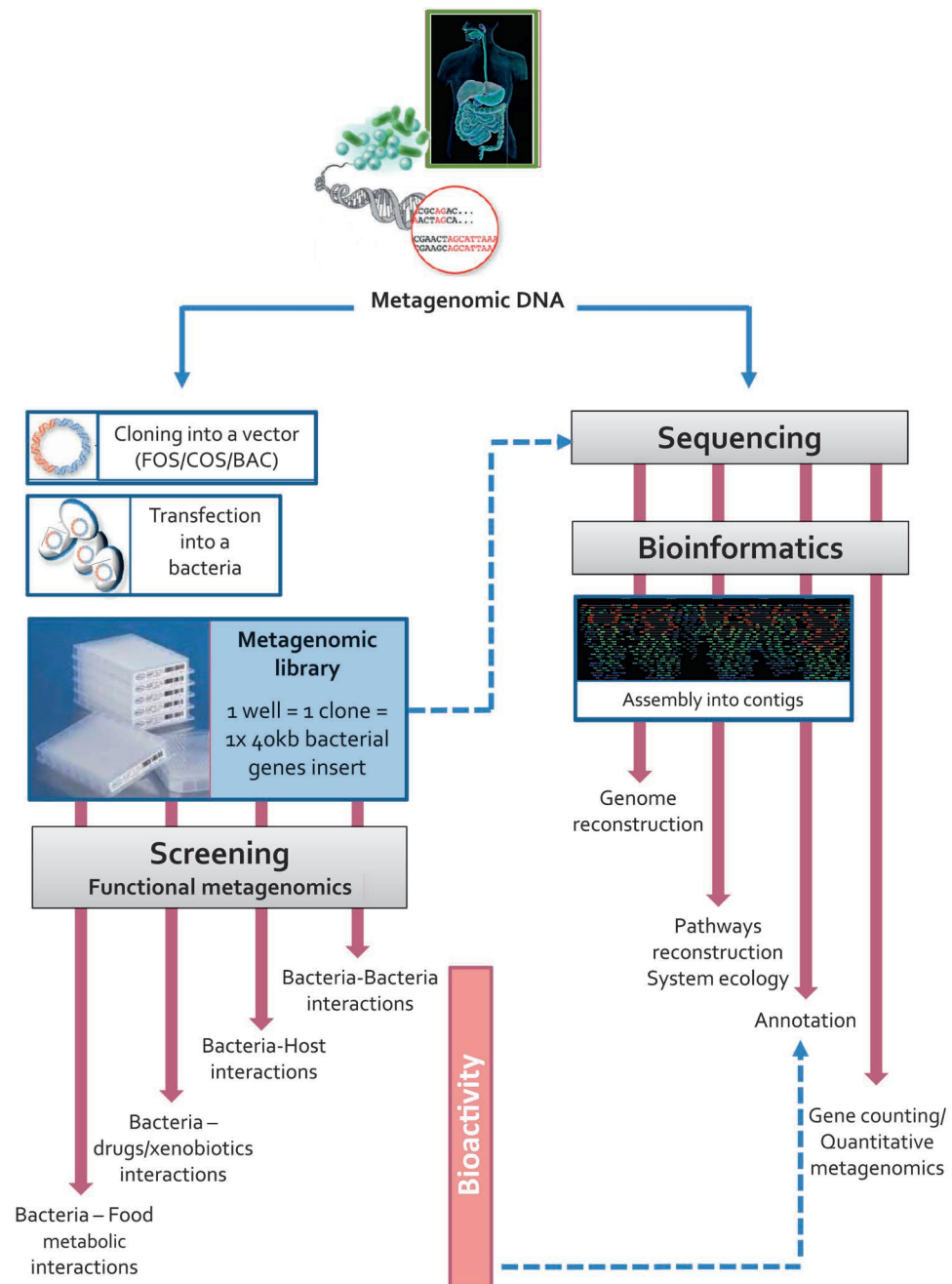


Figure 1 Main metagenomics applications, from the metagenomic libraries construction and screening, until next generation sequencing, gene count and genome reconstruction. Metagenomics libraries are screened for a set of bioactivities ranging from bacteria–food metabolic activities (such as glycoside hydrolases production or butyrate production) to bacteria–bacteria interactions (ie, quorum sensing or antimicrobials activities). These bioactivities are further described by inserts sequencing and annotation. However, metagenomic DNA can be directly sequenced. Following contigs assembly, analyses ranging from bacterial genome reconstruction to systems ecology can be applied. From sequences data, quantitative metagenomics can also be applied to unravel the bacterial genes within an ecosystem. BAC, bacterial artificial chromosome; COS, cosmid; FOS, fosmid.

FUNCTIONAL METAGENOMICS: DISCOVERING NEW FUNCTIONS AND RE-EVALUATING MICROBIAL PATHWAYS

A challenging approach for metagenomic analysis is to identify clones bearing bacterial genes that express a function. Because 86% of bacterial genomic DNA encodes for coding open reading frames,⁴⁴ clones bearing insert fragments of more

than 1 kilobase can be screened for genes and pathways, applying functional metagenomics to the cultured and mostly uncultured bacteria. Although *Escherichia coli* is the most used host for functional metagenomics, other species including *Streptomyces*, *Bacillus subtilis* or *Lactococcus lactis* can be chosen to facilitate the heterologous expression of Gram-positive bacterial DNA.⁴⁵ However, the

taxonomic diversity of bacteria which genes could be expressed in *E coli* has been reported unexpectedly high.⁴⁶

The first screenings of gut metagenomic libraries were applied to the catabolism of dietary fibres, which is of particular interest in human nutrition and health. Dietary fibres have been identified as a strong positive dietary factor in the prevention of obesity, diabetes, colorectal cancer and cardiovascular diseases (World Health Organization, Joint WHO/FAO Expert Consultation "Diet, Nutrition and the Prevention of Chronic Diseases" 2003, WHO Technical Report Series 916). Human gut bacteria produce a vast panel of carbohydrate active enzymes (CAZymes) to degrade components of dietary fibre into metabolisable monosaccharides and disaccharides. CAZymes study was nevertheless restricted to cultivated bacterial species. CAZymes diversity has been described in three metagenomics studies focused on the gut microbiome,^{41 47 48} revealing the presence of at least 81 families of glycoside hydrolases, making the human gut metagenome one of the richest sources of CAZymes.⁴⁹ Tasse *et al* applied high-throughput functional screenings on human gut derived metagenomics clones (156 000 clones representing 5.5 Gb of DNA) and identified new CAZymes.⁵⁰

Similarly, Jones and colleagues screened about 90 000 fosmid clones from the human gut metagenome (representing about 3.6 Gb bacterial DNA) for bile salt hydrolase activity.⁵¹ These functions were found to be present and enriched in all major gut microbial divisions, including archaea. A different strategy allowed the expression of metagenomic β -glucuronidase activity in β -glucuronidase-positive *E coli* fosmid libraries from healthy people and patients with CD.⁵² The study unravelled a new subfamily of β -glucuronidase, dominant in healthy adults and children, which may have specifically evolved to settle in the human gut microenvironment.

Even though the microbiota contributes to the homeostasis of the intestinal mucosa and the maturation of the immune system, the underlying cellular and molecular mechanisms driving the co-evolution and interplay between microbiota and host remain poorly understood. Since the intestinal epithelial cells are the first line in contact with microbes and have a key role in immune regulation and inflammatory pathways, the functional metagenomic approach was also validated for the identification of bacterial genes influencing the proliferation of HT-29 (colorectal carcinoma cell line) human intestinal epithelial cells.⁵³ Using high-throughput screening and reporter gene technology in human epithelial cells, Lakhdari *et al* identified metagenomic clones activating or inhibiting the nuclear factor κ B (NF- κ B) signalling pathway.⁵⁴ Among the bacterial genes modulating the NF- κ B pathway were a putative lipoprotein of yet unknown function and an efflux ABC-type transport system, related to the Lol-D family of lipoprotein transporter complex.

Another application is the screening of plasmid encoding elements. The mobile genetic elements associated with the human gut microbiota (the mobile metagenome) reflect the co-evolution of host and microbe in this community. Genes involved in survival and persistence in the gastrointestinal tract, as well as activities that impact on host-microbe or microbe-microbe interactions, are encoded by the gut mobile metagenome ('mobilome'). Applying the culture independent Transposon-aided capture (TRACA) system, Jones and Marchesi isolated novel plasmids from the human gut microbiota and detected genes enriched in the human gut compared with other ecosystems.⁵⁵ Among the most prevalent functions identified was a putative RelBE toxin-antitoxin addiction module. Moreover, recent evidence from metagenomics indicated that horizontal gene transfer could happen between phylogenetically distant bacterial groups.^{42 50 51} Very recently, Smillie and colleagues decrypted the forces driving these exchanges and demonstrated that ecology is the main driver of gene exchange. For instance, bacteria with the same oxygen tolerance, that is, inhabiting the same body site and sharing ecological features, are actively exchanging genes.⁵⁶

Finally, functional genomics can also provide calibrated assays designed to estimate the functionalities of cultivable microorganisms found to be over-represented or under-represented in several clinical conditions, such as inflammation, or to compare expression of probiotic strains for key functions related to human health.

OUR OTHER GENOME: HIGHLIGHTS OF THE EUROPEAN METAHIT PROJECT

In October 2005, an international meeting gathered together a panel of experts focusing on the challenges and feasibility of an exhaustive analysis of the human metagenome. As an outcome, the International Human Microbiome Consortium (IHMC) was created, the goals of which are to work under a common set of principles and policies to study and understand the role of the human microbiome in the maintenance of health and causation of disease, and to use that knowledge to improve the ability to prevent and treat disease. The Consortium's efforts have focused on generating a shared comprehensive data resource that enables investigators to characterise the relationship between the composition of the human microbiome and human health and disease (<http://www.human-microbiome.org/>). The IHMC's main research programmes are summarised in table 1.

Evidence for co-evolution and the high level of complexity of humans and their microbiome led to the hypothesis of functional redundancy in the intestinal microbiome. Different bacterial species share functional traits and the large amount of intestinal bacteria guarantees the presence of all mandatory functions, providing robustness to this ecosystem. The first extensive catalogue of microbial genes from the human gut, published by Qin

Recent advances in basic science

Table 1 International Human Microbiome Consortium (IHMC) main research programmes

Research programme	Country	Date	Principal investigators' affiliations	Specificities
NIH Jumpstart Program	USA	2007–2008	Baylor College of Medicine, USA; Broad Institute, USA; The J. Craig Venter Institute, USA; Washington University School of Medicine, USA	<ol style="list-style-type: none"> 1. Generation of the complete genome sequences of 200 bacterial strains isolated from the human body 2. Recruitment of donors and securing samples from five body regions 3. Performing 16S rRNA metagenomic sequence analysis of the sampled body regions
MetaHIT, Metagenomics of the Human Intestinal Tract	EU	2008–2011	INRA, France; CEA Genoscope, France; DTU, Denmark; Danone, France; EMBL, Germany; HUVH, Spain; IEO, Italy; NN, Denmark; SDU, Denmark; UCB Pharma, Spain; WU, The Netherlands; WTSI, UK; BGI, People's Republic of China	<ol style="list-style-type: none"> 1. Establishment of an extensive reference catalogue of microbial genes present in the human intestine 2. Development of bioinformatics tools to store, organise and interpret this information. 3. Development of tools to perform quantitative metagenomics 4. Determination of bacterial genes associated with health and disease 5. Development of methods to study the function of bacterial genes associated with disease aiming to understand the underlying mechanisms and host–microbe interactions
MicroObes, Human Intestinal Microbiome in Obesity and Nutritional Transition	France	2008–2010	INRA, France; UPMC, France; CEA Genoscope, France	<ol style="list-style-type: none"> 1. Implementation of concerted nutritional trials in France and China and characterisation of the microbiota of patients using high-throughput metagenomic profiling 2. Sequencing of the dominant gene repertoire of the intestinal microbiota of obese and non-obese volunteers 3. Identification of mechanisms of bacteria–cell crosstalk with relevance in energy harvest from food.
DACC, Data Analysis and Coordination Center	USA	2008–2013	Institute for Genome Sciences, USA	Bacterial strains sequencing (n=354 as of November 2011) and annotation
CMI, Canadian Human Microbiome Initiative	Canada	Start 2010	University of Guelph, Canada; University of British Columbia, Canada; University Health Network, Canada; University of Alberta, Canada; Research Institute of the McGill University Health Centre, Canada; University of Waterloo, Ontario, Canada; University of Calgary, Canada; University of Ottawa, Canada	<ol style="list-style-type: none"> 1. Characterisation of communities of microbes found in different body sites in healthy individuals to determine if individuals share a common microbiome 2. Exploration of the interactions between microbes and determination of how perturbations at one site can affect microbes at another body site 3. Investigation of the role played by the microbiota in human health and probing of the associations between the human microbiome and a variety of diseases and health conditions
The Australian Jumpstart Human Microbiome Project	AU	Start 2009	CSIRO	<ol style="list-style-type: none"> 1. Specific bacterial strains isolations for DAAC sequencing 2. Application of functional metagenomics techniques to unravel the interaction between intestinal microbes and their host 3. Production of practical outcomes for the livestock sector (animal greenhouse gas emissions, reduction of antibiotics use)
HMP, NIH Human Microbiome Project	USA	2008–2012		<ol style="list-style-type: none"> 1. Characterisation of the microbial communities found at several different sites on the human body, including nasal passages, oral cavities, skin, gastrointestinal tract, and urogenital tract 2. Analysis of the role of these microbes in human health and disease 3. Reference set of microbial genome sequences and preliminary characterisation of the human microbiome
Korean Microbiome Diversity, using Korean Twin Cohort Project	Korea	2010–2015	Seoul National University, South Korea	<ol style="list-style-type: none"> 1. Determination of the microbiomes present in various epithelial sites of the human body by using Korean Twin Cohort 2. Investigation of the relationship between human microbiomes and the relevant diseases on their respective sites 3. Development of a database to cover the diversity of Korean microbiomes and their related data 4. Determination of the disease targets for early diagnosis and prevention, and establishment of a dedicated centre for Korean microbiome information and analysis

and colleagues,⁴⁰ described the large variety of traits provided by the intestinal microbiota and an overall metagenome per individual outnumbering by

a factor of 150 the size of the human genome. This catalogue highlighted functions that are important for bacterial survival in the human intestines and

the existence of a functional core, conserved in each individual of the cohort (n=124), despite a high inter-individual specificity of the intestinal microbiota. This functional core or minimal human gut metagenome can be viewed as a set of non-redundant bacterial genes necessary for the normal functioning of the gut ecosystem and encoded across phylogenetic boundaries. In the studied population, 38% of one individual's bacterial genes were shared by at least 50% of the cohort, highlighting a high level of functional similarities between individuals. This percentage fell to 9% when looking at genes shared by at least 80% of the community.

The depth of coverage and wealth of information provided by metagenomic sequencing recently enabled the discovery of distinct 'types' of gut composition in the human population. Through comparison of the phylogenetic and functional composition of the gut microbiota in three different international cohorts of 33, 85 and 166 individuals, Arumugam, Raes and colleagues found that nationality and the presumed similarity in genetic, ethnic and nutritional background was not reflected in gut composition similarity. Instead, they found three distinct clusters of gut composition ('enterotypes'), characterised by the dominant genera and their co-occurring phylogenetic groups that separated the population.⁵⁷ Analysis of the metagenomic data, based solely on the pathways that could be assigned to distinct functions, suggested that enterotypes differ at least in their capacity of vitamin production and metabolic dependencies of the dominant organisms. A recent 16S-based study evaluated the effect of dietary habits of 98 US individuals on their enterotype and found a positive correlation (false discovery rate <25%) between long-term diet (reported habitual

diet over 1 year) and enterotypes. Protein and animal fat uptake was linked to the *Bacteroides*-dominated enterotype while a carbohydrate-rich diet was associated with the *Prevotella*-dominated enterotype. Short-term diet change and 10 days' dietary intervention affected the species composition but not the enterotype distribution.⁵⁸ All in all, these studies still open up a wide range of questions about enterotypes and especially their stability over time in humans as a function of nutrition or clinical status.

FUNCTIONS AND DYSFUNCTIONS OF THE GUT MICROBIOME

The latter studies issued from the MetaHIT programme led to an important breakthrough about functional gut microbiome impairment in inflammation-related diseases, particularly in IBD and obesity. The first study compared 124 individual metagenomes (healthy, overweight and obese individuals, and patients with IBD) and showed that only 25% of one's individual genes are shared with patients with IBD compared with 38% in healthy individuals.⁴⁰ Remarkably, the number of non-redundant bacterial genes was significantly lower in patients with IBD (425 397 ± 126 685, SD; n=25) than in healthy individuals (564 070 ± 121 962, SD; n=99; $p < 10^{-6}$, one-tailed Student t test). Interestingly, functional genes or modules, and also gene counts, strongly correlate with an individual's clinical status (figure 2) and might represent diagnostic tools for numerous human gut disorders. Such a functional microbial dysbiosis in IBD context may affect the host–microbiota crosstalk at the mucosal interfaces (eg, shift in microbial associated molecular patterns exposure repertoire, modification of non-commensal clearance activity or trophic alteration of the global

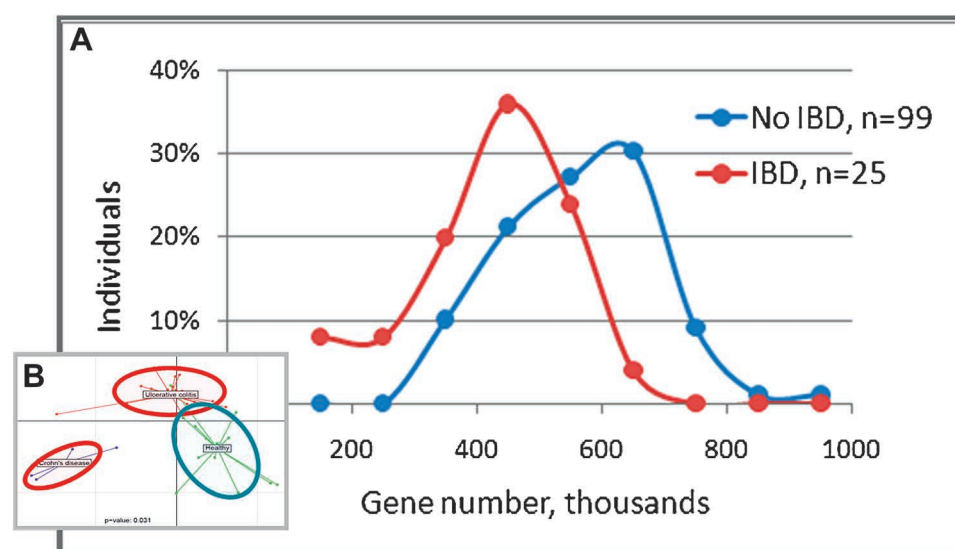


Figure 2 Metagenomic analysis of the human gut microbiota in healthy individuals and patients with inflammatory bowel disease (IBD) highlighted (A) a reduced number of overall non-redundant bacterial genes in patients with IBD. The proportion of individuals having a given number of genes (classes of 100 000 genes were used) is shown. (B) Discrimination of metagenomic composition between healthy individuals (blue) and patients with IBD (red), ulcerative colitis and Crohn's disease, based on an inter-class principal components analysis (adapted from Qin *et al*⁴⁰).

Recent advances in basic science

ecosystem), favouring the onset or at least the maintenance of digestive tract disorders.

In addition, the worldwide epidemic increase in obesity has boosted research on the impact of diet on host health. The so-called western-style high-fat low-fibre diet dramatically impacts the intestinal microbiota. The composition of the intestinal microbiota of lean and obese people also differs, with an elevated abundance of Firmicutes and a lower abundance of Bacteroidetes in people with obesity that can be altered with diet-induced weightloss.²⁸ Turnbaugh *et al* have pointed out the existence of a core microbiome between lean and obese individuals while compositional analysis indicated a microbial dysbiosis affecting the Firmicutes/Bacteroidetes ratio.⁴⁷

Other implications of the composition and functioning of the intestinal microbiota are less intuitive as they seem to trespass physiological boundaries. In humans, the role of the intestinal microbiota in autistic spectrum disorders, for example, remains speculative. Many autistic children are also reported to have gastrointestinal problems,⁵⁹ which again raises the question of a causal link between events or simply co-occurring events. However, when assessing the effects of intestinal dysbiosis, the

gut–brain axis and its bidirectional nature should no longer be ignored.^{60–61} Several animal studies recently challenged the concept of free will and demonstrated that behaviour (anxiety, mating preference, eating behaviour etc) can be modified by intestinal microbiota modulation.^{62–64}

The potential clinical importance of enterotypes distribution and stratification is not to be underestimated: currently, multiple clinical studies are investigating associations between enterotypes and various pathologies such as obesity, diabetes and IBD. If clear links are found, this will open the way to enterotype-based diagnostic and prognostic tests. In addition, enterotypes might be associated with different responses to treatment and even differences in drug metabolism, which may require enterotype-based patient stratification strategies in care management and treatment of various illnesses.

TOWARDS AN INTEGRATED MICROBIOMIC VISION OF THE HUMAN BODY

Metagenomics is a very powerful tool allowing the description of the genetic potential of the micro-organisms present in a defined environment, but does not permit monitoring of their activity or gene

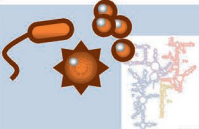
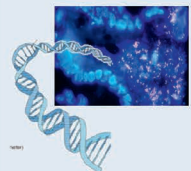
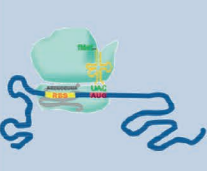

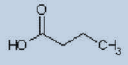
Meta-omics		Molecule	Knowledge	Limits	Clinical implications
Phylogeny		16S rDNA	► Bacterial composition & diversity	► No information on bacterial functions ► Except archaea	► Composition dysbiosis ► Healthy or disease specific species
Metagenomics		Chromosomal genomic DNA	► High resolution Microbiota profiling ► Genes contents from uncultivated microbes	► No information on microbial expressed functions	► Functional dysbiosis ► Healthy or disease specific microbial genes ► Toward diagnostics functional based studies
Metatranscriptomics		Messenger RNA/ cDNA	► High resolution gene expression profiling ► Differential microbial gene expression various physiological/ environmental conditions	► Poor stability of bacterial mRNA ► Representativity unknown / Multiple purification steps needed ► No unique protocol	► Functional Dysbiosis ► Microbial activity kinetics ► Expressed genes at specific time and location ► Specific monitoring of active bacteria
Metaproteomics		Proteins/ Peptides	► High resolution protein mapping and profiling ► Differential microbial proteins production under various physiological/ environmental conditions	► Many unknown proteins in databases ► Heterogeneous stability ► No unique protocol	► Function confirmed: genome annotation improvement ► Eucaryotes-procaryotes analogs identification ► Biomarkers
Metabolomics		Metabolites	► Microbial and host Metabolic profiling ► Easy to perform on very low amount of material feces/serum/urine	► Many unknown metabolites in databases ► Strict Identification of compound tedious ► No unique protocol ► Combination of host and microbial molecules	► New pathways confirmed or identified ► High throughput metabolomic screening of biomarkers ► Easy translation to clinical setting

Figure 3 Integrated meta-omics. The different levels of analyses are represented from phylogeny to metabolomics.

expression. After assessing the functional capacity of the intestinal microbiota, the next level is to assess the actual functions exerted by the microbiota (figure 3). One way to do so is by analysing the mRNA sequences from the microbiota, the metatranscriptome. Frias-Lopez *et al* have applied metagenomics and metatranscriptomics to analyse a marine ecosystem.⁶⁵ They reported that microbial presence and activity are not linked as the most expressed genes were not the most represented ones within the metagenome. They concluded that understanding the fine structure, dynamics and overall impact of a microbial ecosystem requires the combination of these two approaches. Metatranscriptomics analysis of faecal microbiomes from 10 healthy humans⁶⁶ showed that carbohydrate metabolism, energy production and synthesis of cellular components were the main expressed functionalities. In contrast, housekeeping activities such as amino acid and lipid metabolism were lowered in the metatranscriptome. However, bacterial DNA/RNA extraction methods have an important influence on the downstream results.⁶⁷ The specific impact of different sampling and nucleic acid extraction methods on metagenomics, and also meta-omics data is currently assessed within the IHMS European funded project International Human Microbiome Standards (IHMS; <http://www.microbiome-standards.org/>). The main goal of the IHMS is to harmonise practices and facilitate data comparison among projects worldwide.

Because mRNA stability is low, microbial gene expression is better captured if sampled physically close to the expression site. However, to compare the transcriptome of bacteria at different intestinal sites (luminal, adherent, mucosal), a protocol uniformly efficient on different types of material (faeces, biopsies, mucus) is mandatory and to our knowledge no satisfactory protocol for bacterial RNA extraction from each of these different matrices has been published so far.

Knowing exactly which proteins and metabolites are active will broaden our knowledge of the intestinal ecosystem. Future challenges for a better understanding of human biology would be clearly related to an integrative perception of our gut microbiome obtained from meta-omics application. In a metaproteomic study of two healthy individuals,⁶⁸ comparison with metagenomic data drew an interesting picture showing an overabundance of proteins related to post-translational modifications, protein folding and turnover. In contrast, levels of proteins involved in inorganic ion metabolism, cell wall and membrane biogenesis, cell division and secondary metabolite biosynthesis were lower. The data integration process of metagenomic and metaproteomic outcomes has already provided new insights into the genetic capabilities and usage of the human gut microbiome. Meta-metabolomics is currently the most widely used meta-omics approach to study the human gut microbiome. Mainly applied to faecal water extracts, metabolomics has already been used to study numerous

digestive tract disorders, including colorectal cancer, IBD and IBS. Described as a suitable approach to be combined with metagenomics,⁶⁹ it aims to assess the metabolite catalogues of a biome according to a given physiological and environmental context. Its application to healthy and IBD faecal extracts has shed light on metabolome differences among healthy individuals, patients with CD and patients with ulcerative colitis (UC).⁷⁰ Consistent with previous phylogenetic studies, the IBD metabolome was characterised by reduced levels of butyrate, acetate, and also methylamine and trimethylamine, and an elevated quantity of amino acids. The authors concluded that this was reminiscent of an IBD microbiome composition shift or dysbiosis affecting both short-chain fatty acid production and nutrient absorption. More specifically, the CD metabolome was associated with more extensive inflammation compared with the UC metabolome. Recently, metabolomics was also applied to highlight biomarker metabolites of the CD microbiome.⁷¹ Metabolites issued from the metabolism of amino acids, fatty acids especially arachidonic acids, and bile acids were the most notably affected in CD. Another study, including patients with UC and IBS, described an increase in taurine and cadaverine quantities in UC while patients with IBS had higher bile acid concentration and lower levels of branched chain fatty acids compared with relative controls.⁷² However, changes in short-chain fatty acid and amino acid concentrations were not significant. Finally, metabolomes of patients with colorectal cancer were also investigated and quantities of short-chain fatty acids (butyrate and acetate) and proline/cysteine were respectively lower or higher compared with healthy metabolomes.

THE FORGOTTEN MICROORGANISMS OF INTESTINAL MICROBIOTA

In almost all investigated ecosystems, there are around 10 phages for every microbial cell, making phages the most abundant biological entities on the planet. Human gut metagenomics highlighted high rates of prophages (ie, bacteriophage genes inserted in the bacterial chromosomes) within bacterial metagenomes. By killing microorganisms, phages greatly influence global biogeochemical cycles, and have been predicted to help maintain microbial species diversity. In 2003, using metagenomics, Breitbart *et al* described over 1300 viral genotypes within one human faecal sample, most of them corresponding to unknown bacteriophages.⁷³ In addition, Reyes *et al* demonstrated that faecal viromes, studied in twins and their mothers, were unique to individuals regardless of their degree of genetic relatedness.⁷⁴ Very recently, Minot *et al* confirmed these data and reported that viromes were highly sensitive to diet changes.⁷⁵

More strikingly, much higher amounts of bacteriophages were described in the mucosa of healthy individuals and patients with CD.⁷⁶ Virus-like particles were counted up to 4×10^9 per biopsy, and

Siphoviridae, Myoviridae and Podoviridae families' morphotypes were dominant. In the context of CD, the ulcerated mucosa exhibited markedly less free bacteriophages than the non-ulcerated mucosa. How the bacteria/phages ratio compares in these very distinct regions of the CD mucosa is a critical question. The current working paradigm for microbial–viral community dynamics is the Lotka Volterra model. Colloquially known as Kill-the-Winner, it predicts that viruses will rapidly and drastically reduce the population of the most abundant microbial species, preventing the best microbial competitors from building up a highly dominant biomass. However, Paterson *et al* recently validated the Red Queen theory.⁷⁷ Named after Lewis Carroll's character from *Through the Looking Glass*, in which the Queen tells Alice: 'It takes all the running you can do to keep in the same place', this theory hypothesises that bacteria and viruses are in a constant evolutionary arms race. Each has to evolve ever-better ways of outwitting the other to avoid losing out. In addition, these important biomasses of bacteria and phages at the mucosal side are in the vicinity of the toll-like receptors (TLRs) and the immune system. A role for bacteriophages in IBD pathogenesis has been suspected for years and several mechanisms have been hypothesised.⁷⁸ Upon environmental changes, phages can notably quickly go from a lysogenic into a lytic phase and kill bacteria, which could, in turn,

locally decrease specific bacterial populations herein dynamically changing the eubiosis or dysbiosis. Phages might be involved in IBD pathogenesis by changing the recognition patterns of bacteria. By covering bacterial cells, they would change the detection of microbial motifs by TLRs. Finally, phages could be directly recognised by the host and trigger specific modulations. Viruses are known to be recognised by TLR3⁷⁹ and TLR13,⁸⁰ but so far no TLR recognising specific phage components has been described.

Interestingly, the first bacteriophages were isolated from stools of patients recovering from dysenteric diseases.⁸¹ D'Herelle rapidly demonstrated that bacteriophages could be used to treat enteric bacterial infections, creating a new field called 'phage therapy'.^{82 83} Even though their usage in medicine has been controversial, the use of bacteriophages in medicine in east Europe was further developed after the Second World War.^{84 85} Oral administration to humans was shown to be safe⁸⁶ and nowadays a renewed interest in phage therapy is growing. The first challenge in applying phage therapy in humans is the knowledge of phage host. The specificity of phages is also critical because they could be used as cocktails for larger spectra. Furthermore, the dosage, lethality or survival of phages, isolation and preparation for commercial phage therapy need to be carefully characterised.⁸⁷

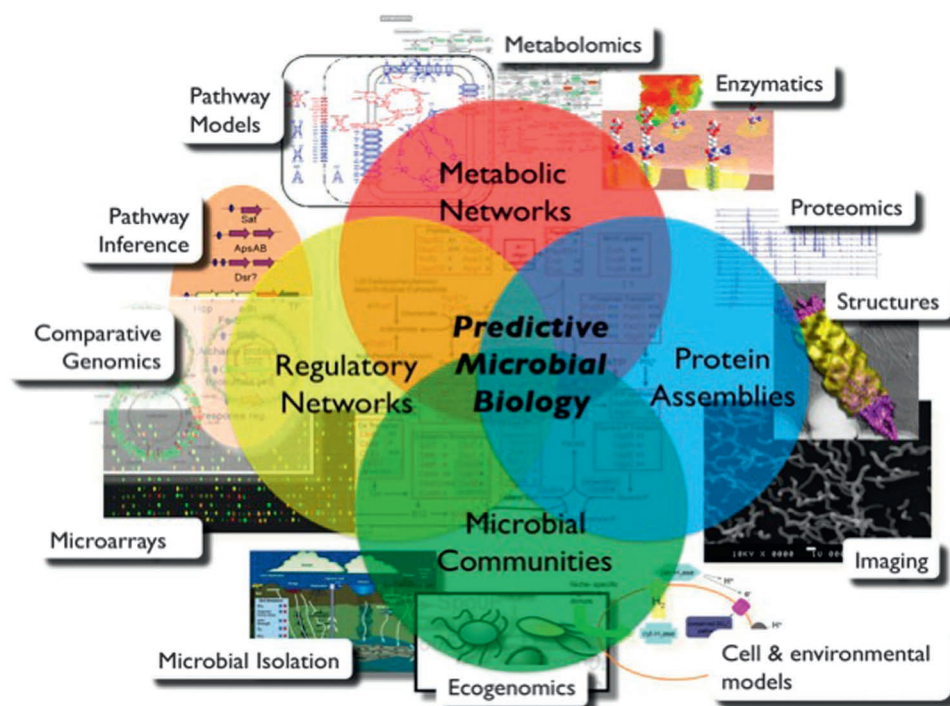


Figure 4 An integrative approach to study microbial systems. Modified from <http://ms.lbl.gov/>, courtesy of T Northern. The comprehension of a complex microbial system and the prediction of its dynamic require the integration of heterogeneous data obtained using several of the 12 methods depicted in the rectangles. For example, microbial isolation and metagenomics allow characterization of the microbial communities, and they can now be combined with metabolomics and pathway models to obtain the metabolic network (of the organisms or) of the whole community. In addition, regulation of these microbial systems, critical in health and disease, can be defined using in silico comparative genomics, to infer pathways that might be further validated against microarray data.

Glossary

Core human microbiome

Whatever is shared in a given body habitat among all or the vast majority of human microbiomes. A core microbiome may include a common set of organisms (phylotypes or their genomes), gene or protein families, and/or metabolic capabilities. Microbial genes that are variably represented in different humans may contribute to our distinctive metabolic attributes (metabotypes).

Dysbiosis

Alteration of the microbiota in comparison to the normal, healthy state. Dysbiosis refers to a condition of microbial imbalances compared with the 'eubiosis' condition, a state of equilibrium between bacterial symbionts and pathobionts. Dysbiosis is mostly reported in the digestive tract or on the skin, but can also occur on any exposed surface or mucous membrane. Dysbiosis can be observed at several analytical levels, from diversity to compositional or functional imbalances.

Meta-omics

Meta-omics includes shotgun sequencing of microbial DNA isolated directly from a given environment; high-throughput screening of expression libraries, constructed from cloned community DNA, to identify specific functions such as antibiotic resistance (functional metagenomics); profiling of RNAs and proteins produced by a microbiome (meta-transcriptomics and meta-proteomics); and identification of a community's metabolic network (metabolomics).

Metabolomics

The characterisation by mass spectroscopy, nuclear magnetic resonance, or other analytical methods of metabolites generated by one or more organisms in a given physiological and environmental context.

Metagenomics

The study of metagenomes, genetic material recovered directly from environmental samples. Metagenomics is the genomic analysis (analysis of all the DNA in an organism) applied to all the microorganisms of a microbial ecosystem without previous identification. Metagenomics is an emerging field encompassing culture-independent studies of the structures and functions of microbial communities and their interactions with the habitats they occupy to understand their biological diversity.

Microbiota

A microbial community, including bacteria, archaea, eukarya, and viruses, which occupies a given habitat.

Microbiome

The totality of microbes, their genetic elements (genomes), and environmental interactions in a defined environment. In this sense, the human 'micro'biome would be defined as the collection of microorganisms associated with the human body, and their collective genomes would constitute a metagenome. However, the term microbiome is now commonly used to refer to the collective genomes present in members of a given microbiota.

Open reading frames

DNA sequence that does not contain a stop codon in a given reading frame. It allows initial identification of candidate protein coding regions in a DNA sequence. A simple gene prediction algorithm for prokaryotes might look for a start codon followed by an open reading frame that is long enough to encode a typical protein, where the codon usage of that region matches the frequency characteristic for the given organism's coding regions.

Phylotype

A phylogenetic group of microbes, currently defined by a threshold per cent identity shared among their small subunit (16S) rRNA genes (eg, $\geq 97\%$ for a 'species' level phylotype).

THE HUMAN GUT MICROBIOME: UNSOLVED PROBLEMS, NEW DIRECTIONS AND EMERGING CHALLENGES

The potential to develop new microbial diagnostic markers of the human metabolic status, to provide early disease diagnostic techniques and new therapeutic strategies, or to maximise the contribution of our microorganisms or probiotics has spawned many human microbiome projects. However, until now, functions encoded by numerous microbial genes found predominant in healthy individuals while missing in patients with disease have so far not been

identified. The ongoing collaboration within the IHMC allows metagenomic projects to be supported by strong datasets originating from ongoing genomes sequencing through human microbiome projects. These new experimental and computational tools will facilitate the discovery of impaired pathways to be specifically targeted at larger cohorts using omics technologies.

Biological validation on large human cohorts is required to demonstrate the robustness of predictors of diseases. The classification of human gut microbiome into three enterotypes opens up

Recent advances in basic science

Web resources

- ▶ **IHMC:** <http://www.human-microbiome.org/>. The International Human Microbiome Consortium focuses on generating a shared comprehensive data resource that enables investigators to characterise the relationship between the composition of the human microbiome and human health and disease.
- ▶ **MetaHit:** <http://www.metahit.eu/>. MetaHit is a project financed by the European Commission under the 7th Framework Program and coordinated by S D Ehrlich, INRA, France. The consortium gathers 13 partners from academia and industry, from a total of eight countries.
- ▶ **HMP:** <http://commonfund.nih.gov/hmp/overview.aspx>. The aim of the Human Microbiome Program, funded by the NIH in the USA, is to characterise microbial communities found at multiple human body sites and to look for correlations between changes in the microbiome and human health.
- ▶ **Gold:** <http://www.genomesonline.org/cgi-bin/index.cgi>. Genomes Online Database, is a web resource for comprehensive access to information regarding genome and metagenome sequencing projects, and their associated metadata, around the world.

questions about their functional impact on human health and individual 'classifications'. Even though larger studies analysing more diverse cohorts will be necessary to confirm and possibly refine the enterotype concept, whether the enterotypes correspond to ecosystems with different metabolic properties, but also distinct community–host interactions, is a key question. The enterotypes stability, or 'enterostates' variability, also needs to be investigated because it will condition their use in nutritional and clinical trials. The original enterotype work had very few family members; although some of these were in different enterotypes, a (partial) genetic/immunological/maternal effect should not be ruled out. Also, current studies have focused on limited and western-diet biased cohorts. A broader sampling of the human population, using bigger cohorts⁸⁸ with multiple age groups, wider

geographical spread including remote populations, and for longer periods of time will increase our understanding of the true nature of enterotypes and pinpoint their underlying origins.

If faecal sample analysis as a surrogate for the entire gut microbiota is necessary for large cohort screening and specific biomarker detection, determining the mechanisms will require the metagenomics description of the mucosal microbiota. Recently, Wang *et al* showed that, from the metagenome associated with the right and left colonic mucosa,⁸⁹ functional diversity was comparable to that reported in faeces, with an overall functionality mainly supported by genes associated with carbohydrate, protein and nucleic acid utilisation. At mucosal interfaces, a loss of bacteria–host crosstalk was described in patients with UC by combining a description of the microbiota and host transcriptomic profiles.²⁶ The co-sequencing and direct interfacing of human and microbiome genomes is now feasible and, while computationally challenging, it should uncover a new link between human–microbes co-evolution and pathways disturbance in disease conditions.

More and more evidence from animal studies shows a genetic effect on microbiome composition (especially with regard to innate immunity-related mutations).^{90–92} Interestingly, unaffected relatives of patients with IBD also show subclinical microbial dysbiosis.^{26 93 94} However, the genetic and environmental risk factors they share with their diseased relatives seem to be counterbalanced. For example, unaffected relatives of patients with IBD have more butyrate-producing bacteria. Therefore, metagenomic analysis of unaffected relatives will be important to answer the recurrent question of whether the dysbiosis is the 'cause or consequence' of the disease.

Gnotobiotic animal models, colonised by synthetic enterotypes microbiota, could represent suitable tools for translational research to characterise the impact of genetic defects, diet alteration, probiotic intake and microbiome transplant, and to provide the proof of principle needed to direct and interpret human studies.

Finally, metagenomics has strongly increased our capacity to mine the clinical and biological data that will soon allow construction and experimental validation of biological models. Mathematical models of the human gut, first focusing on host–microbiota interaction or on specific microbial pathways, should soon be integrated in a global model framework capturing key inter-relationships between host factors (genetic, epigenetic, immunological and serological data) and its microbiome (figure 4). Since many diseases are not spatially constrained to one organ, as discussed at the October 2005 IHMC foundation meeting, the understanding of human–microbiota interactions requires the analysis of all human colonised body site metagenomes. This will lead to the development of novel tools for preclinical prognostics and early diagnostics; definition and validation of biological networks relevant to different

Key messages

- ▶ The human gut contains 10^{14} bacteria (mass of 2 kg in the gut), 10 times the number of human cells, 150 times more genes than the human genome.
- ▶ Other organisms in the human gut belong to archaea, viruses (prokaryotic and eukaryotic), parasites, fungi.
- ▶ The human gut microbiota plays a key role in human health, gastrointestinal disease and other diseases.
- ▶ Commensal bacteria develop mechanisms of crosstalk with human cells that play key roles in homeostasis maintenance.
- ▶ Metagenomics allows the description of the combined genomes of the microorganisms present in the gut, even non-cultured ones, giving access to their potential functions.
- ▶ Human gut contains an important diversity and high number of non-redundant bacterial genes.
- ▶ Based on gut microbial composition distribution, the human population can be subdivided into enterotypes.
- ▶ A better description of the human gut microbiota will lead to description of biomarkers of disease, development of new probiotics/prebiotics, new therapies such as faecal transplantation.

disease pathophysiologies; and identification of new therapeutic targets and intervention/prevention strategies.

Acknowledgements The authors would like to thank Nicolas Lapaque for careful proofreading of the manuscript.

Contributors All authors have contributed to and approved the manuscript.

Funding The authors would like to gratefully acknowledge the financial support of the French National Agency for Research, the 7th European Framework Program (MetaHIT project and Crosstalk project). JR is supported by the Research Foundation Flanders (FWO) and the Agency for Innovation by Science and Technology (IWT).

Competing interests None.

Provenance and peer review Commissioned; internally peer reviewed.

Data sharing statement Available data in this paper will be shared upon request to the corresponding author.

REFERENCES

- Davies J. In a map for human life, count the microbes, too. *Science* 2001;**291**:2316.
- Berg RD. The indigenous gastrointestinal microflora. *Trends Microbiol* 1996;**4**:430–5.
- Savage DC. Microorganisms associated with epithelial surfaces and stability of the indigenous gastrointestinal microflora. *Nahrung* 1987;**31**:383–95.
- Edwards DP. The roles of tolerance in the evolution, maintenance and breakdown of mutualism. *Naturwissenschaften* 2009;**96**:1137–45.
- Gaboriau-Routhiau V, Rakotobe S, Lecuyer E, et al. The key role of segmented filamentous bacteria in the coordinated maturation of gut helper T cell responses. *Immunity* 2009;**31**:677–89.
- Neish AS. Microbes in gastrointestinal health and disease. *Gastroenterology* 2009;**136**:65–80.
- Round JL, Mazmanian SK. The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol* 2009;**9**:313–23.
- Bocci V. The neglected organ: bacterial flora has a crucial immunostimulatory role. *Perspect Biol Med* 1992;**35**:251–60.
- Backhed F, Ley RE, Sonnenburg JL, et al. Host–bacterial mutualism in the human intestine. *Science* 2005;**307**:1915–20.
- Cummings JH. Microbial digestion of complex carbohydrates in man. *Proc Nutr Soc* 1984;**43**:35–44.
- Hooper LV, Midtvedt T, Gordon JL. How host–microbial interactions shape the nutrient environment of the mammalian intestine. *Annu Rev Nutr* 2002;**22**:283–307.
- Kelly D, Campbell JI, King TP, et al. Commensal anaerobic gut bacteria attenuate inflammation by regulating nuclear-cytoplasmic shuttling of PPAR-gamma and RelA. *Nat Immunol* 2004;**5**:104–12.
- Willing BP, Van Kessel AG. Enterocyte proliferation and apoptosis in the caudal small intestine is influenced by the composition of colonizing commensal bacteria in the neonatal gnotobiotic pig. *J Anim Sci* 2007;**85**:3256–66.
- Suau A, Bonnet R, Sutren M, et al. Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. *Appl Environ Microbiol* 1999;**65**:4799–807.
- Zoetendal EG, Rajilic-Stojanovic M, de Vos WM. High-throughput diversity and functionality analysis of the gastrointestinal tract microbiota. *Gut* 2008;**57**:1605–15.
- Koenig JE, Spor A, Scalfone N, et al. Succession of microbial consortia in the developing infant gut microbiome. *Proc Natl Acad Sci U S A* 2011;**108**(Suppl 1):4578–85.
- Mackie RI, Sghir A, Gaskins HR. Developmental microbial ecology of the neonatal gastrointestinal tract. *Am J Clin Nutr* 1999;**69**:1035S–45S.
- De La Cochetiere MF, Durand T, Lepage P, et al. Resilience of the dominant human fecal microbiota upon short-course antibiotic challenge. *J Clin Microbiol* 2005;**43**:5588–92.
- Dethlefsen L, Relman DA. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. *Proc Natl Acad Sci U S A* 2011;**108**(Suppl 1):4554–61.
- Frank DN, St Amand AL, Feldman RA, et al. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci U S A* 2007;**104**:13780–5.
- Tap J, Mondot S, Levenez F, et al. Towards the human intestinal microbiota phylogenetic core. *Environ Microbiol* 2009;**11**:2574–84.
- Dethlefsen L, Eckburg PB, Bik EM, et al. Assembly of the human intestinal microbiota. *Trends Ecol Evol* 2006;**21**:517–23.
- Ley RE, Hamady M, Lozupone C, et al. Evolution of mammals and their gut microbes. *Science* 2008;**320**:1647–51.
- Backhed F, Ding H, Wang T, et al. The gut microbiota as an environmental factor that regulates fat storage. *Proc Natl Acad Sci U S A* 2004;**101**:15718–23.
- Frank DN, Robertson CE, Hamm CM, et al. Disease phenotype and genotype are associated with shifts in intestinal-associated microbiota in inflammatory bowel diseases. *Inflamm Bowel Dis* 2011;**17**:179–84.
- Lepage P, Hasler R, Spehlmann ME, et al. Twin study indicates loss of interaction between microbiota and mucosa of patients with ulcerative colitis. *Gastroenterology* 2011;**141**:227–36.
- Ley RE, Backhed F, Turnbaugh P, et al. Obesity alters gut microbial ecology. *Proc Natl Acad Sci U S A* 2005;**102**:11070–5.
- Ley RE, Turnbaugh PJ, Klein S, et al. Microbial ecology: human gut microbes associated with obesity. *Nature* 2006;**444**:1022–3.
- Manichanh C, Rigottier-Gois L, Bonnaud E, et al. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. *Gut* 2006;**55**:205–11.
- Mondot S, Kang S, Furet JP, et al. Highlighting new phylogenetic specificities of Crohn's disease microbiota. *Inflamm Bowel Dis* 2011;**17**:185–92.
- Seksik P, Sokol H, Lepage P, et al. Review article: the role of bacteria in onset and perpetuation of inflammatory bowel disease. *Aliment Pharmacol Ther* 2006;**24**(Suppl 3):11–18.
- Sartor RB. Intestinal microflora in human and experimental inflammatory bowel disease. *Curr Opin Gastroenterol* 2001;**17**:324–30.
- Nagalingam NA, Lynch SV. Role of the microbiota in inflammatory bowel diseases. *Inflamm Bowel Dis*. Published Online First: 20 September 2011. doi:10.1002/ibd.21866
- Tamboli CP, Neut C, Desreumaux P, et al. Dysbiosis as a prerequisite for IBD. *Gut* 2004;**53**:1057.
- Handelsman J, Rondon MR, Brady SF, et al. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem Biol* 1998;**5**:R245–9.
- Rondon MR, Raffel SJ, Goodman RM, et al. Toward functional genomics in bacteria: analysis of gene expression in *Escherichia coli* from a bacterial artificial chromosome library of *Bacillus cereus*. *Proc Natl Acad Sci U S A* 1999;**96**:6451–5.
- Niedringhaus TP, Milanova D, Kerby MB, et al. Landscape of next-generation sequencing technologies. *Anal Chem* 2011;**83**:4327–41.
- Pope PB, Smith W, Denman SE, et al. Isolation of *Succinivibrionaceae* implicated in low methane emissions from Tamar wallabies. *Science* 2011;**333**:646–8.
- Tyson GW, Chapman J, Hugenholtz P, et al. Community structure and metabolism through reconstruction of microbial genomes from the environment. *Nature* 2004;**428**:37–43.
- Qin J, Li R, Raes J, et al. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 2010;**464**:59–65.
- Gill SR, Pop M, Deboy RT, et al. Metagenomic analysis of the human distal gut microbiome. *Science* 2006;**312**:1355–9.
- Kurokawa K, Itoh T, Kuwahara T, et al. Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. *DNA Res* 2007;**14**:169–81.
- Manichanh C, Chapelle CE, Frangeul L, et al. A comparison of random sequence reads versus 16S rDNA sequences for estimating the biodiversity of a metagenomic library. *Nucleic Acids Res* 2008;**36**:5180–8.
- Davenport CF, Tummler B. Abundant oligonucleotides common to most bacteria. *PLoS One* 2010;**5**:e9841.
- Gabor EM, Alkema WB, Janssen DB. Quantifying the accessibility of the metagenome by random expression cloning techniques. *Environ Microbiol* 2004;**6**:879–86.
- Handelsman J. Metagenomics: application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* 2004;**68**:669–85.
- Turnbaugh PJ, Hamady M, Yatsunenko T, et al. A core gut microbiome in obese and lean twins. *Nature* 2009;**457**:480–4.

48. **Turnbaugh PJ**, Henrichs B, Gordon JL. Viewing the human microbiome through three-dimensional glasses: integrating structural and functional studies to better define the properties of myriad carbohydrate-active enzymes. *Acta Crystallogr Sect F Struct Biol Cryst Commun* 2010;**66**:1261–4.
49. **Li LL**, McCorkle SR, Monchy S, *et al*. Bioprospecting metagenomes: glycosyl hydrolases for converting biomass. *Biotechnol Biofuels* 2009;**2**:10.
50. **Tasse L**, Bercovici J, Pizzut-Serin S, *et al*. Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes. *Genome Res* 2010;**20**:1605–12.
51. **Jones BV**, Begley M, Hill C, *et al*. Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome. *Proc Natl Acad Sci U S A* 2008;**105**:13580–5.
52. **Gloux K**, Berteau O, El Oumami H, *et al*. A metagenomic beta-glucuronidase uncovers a core adaptive function of the human intestinal microbiome. *Proc Natl Acad Sci U S A* 2011;**108**(Suppl 1):4539–46.
53. **Gloux K**, Leclerc M, Ilizier H, *et al*. Development of high-throughput phenotyping of metagenomic clones from the human gut microbiome for modulation of eukaryotic cell growth. *Appl Environ Microbiol* 2007;**73**:3734–7.
54. **Lakhdari O**, Cultrone A, Tap J, *et al*. Functional metagenomics: a high throughput screening method to decipher microbiota-driven NF-kappaB modulation in the human gut. *PLoS One* 2010;**5**:pii: e13092.
55. **Jones BV**, Marchesi JR. Transposon-aided capture (TRACA) of plasmids resident in the human gut mobile metagenome. *Nat Methods* 2007;**4**:55–61.
56. **Smillie CS**, Smith MB, Friedman J, *et al*. Ecology drives a global network of gene exchange connecting the human microbiome. *Nature* 2011;**480**:241–4.
57. **Arumugam M**, Raes J, Pelletier E, *et al*. Enterotypes of the human gut microbiome. *Nature* 2011;**473**:174–80.
58. **Wu GD**, Chen J, Hoffmann C, *et al*. Linking long-term dietary patterns with gut microbial enterotypes. *Science* 2011;**334**:105–8.
59. **Horvath K**, Papadimitriou JC, Rabsztyan A, *et al*. Gastrointestinal abnormalities in children with autistic disorder. *J Pediatr* 1999;**135**:559–63.
60. **Collins SM**, Bercik P. The relationship between intestinal microbiota and the central nervous system in normal gastrointestinal function and disease. *Gastroenterology* 2009;**136**:2003–14.
61. **Gonzalez A**, Stombaugh J, Lozupone C, *et al*. The mind–body–microbial continuum. *Dialogues Clin Neurosci* 2011;**13**:55–62.
62. **Heijtz RD**, Wang S, Anuar F, *et al*. Normal gut microbiota modulates brain development and behavior. *Proc Natl Acad Sci U S A* 2011;**108**:3047–52.
63. **Neufeld KM**, Kang N, Bienenstock J, *et al*. Reduced anxiety-like behavior and central neurochemical change in germ-free mice. *Neurogastroenterol Motil* 2011;**23**:255–64, e119.
64. **Sharon G**, Segal D, Ringo JM, *et al*. Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* 2010;**107**:20051–6.
65. **Frias-Lopez J**, Shi Y, Tyson GW, *et al*. Microbial community gene expression in ocean surface waters. *Proc Natl Acad Sci U S A* 2008;**105**:3805–10.
66. **Gosalbes MJ**, Durban A, Pignatelli M, *et al*. Metatranscriptomic approach to analyze the functional human gut microbiota. *PLoS One* 2011;**6**:e17447.
67. **O Cuiv P**, Aguirre de Carcer D, Jones M, *et al*. The effects from DNA extraction methods on the evaluation of microbial diversity associated with human colonic tissue. *Microb Ecol* 2011;**61**:353–62.
68. **Verberkmoes NC**, Russell AL, Shah M, *et al*. Shotgun metaproteomics of the human distal gut microbiota. *ISME J* 2009;**3**:179–89.
69. **Turnbaugh PJ**, Gordon JL. An invitation to the marriage of metagenomics and metabolomics. *Cell* 2008;**134**:708–13.
70. **Marchesi JR**, Holmes E, Khan F, *et al*. Rapid and noninvasive metabonomic characterization of inflammatory bowel disease. *J Proteome Res* 2007;**6**:546–51.
71. **Jansson J**, Willing B, Lucio M, *et al*. Metabolomics reveals metabolic biomarkers of Crohn's disease. *PLoS One* 2009;**4**:e6386.
72. **Le Gall G**, Noor SO, Ridgway K, *et al*. Metabolomics of fecal extracts detects altered metabolic activity of gut microbiota in ulcerative colitis and irritable bowel syndrome. *J Proteome Res* 2011;**10**:4208–18.
73. **Breitbar M**, Hewson I, Felts B, *et al*. Metagenomic analyses of an uncultured viral community from human feces. *J Bacteriol* 2003;**185**:6220–3.
74. **Reyes A**, Haynes M, Hanson N, *et al*. Viruses in the faecal microbiota of monozygotic twins and their mothers. *Nature* 2010;**466**:334–8.
75. **Minot S**, Sinha R, Chen J, *et al*. The human gut virome: inter-individual variation and dynamic response to diet. *Genome Res* 2011;**21**:1616–25.
76. **Lepage P**, Colombet J, Marteau P, *et al*. Dysbiosis in inflammatory bowel disease: a role for bacteriophages? *Gut* 2008;**57**:424–5.
77. **Patterson S**, Vogwill T, Buckling A, *et al*. Antagonistic coevolution accelerates molecular evolution. *Nature* 2010;**464**:275–8.
78. **Riley PA**. Bacteriophages in autoimmune disease and other inflammatory conditions. *Med Hypotheses* 2004;**62**:493–8.
79. **Alexopoulou L**, Holt AC, Medzhitov R, *et al*. Recognition of double-stranded RNA and activation of NF-kappaB by toll-like receptor 3. *Nature* 2001;**413**:732–8.
80. **Shi Z**, Cai Z, Sanchez A, *et al*. A novel toll-like receptor that recognizes vesicular stomatitis virus. *J Biol Chem* 2011;**286**:4517–24.
81. **D'Herelle F**. On an invisible microbe antagonistic toward dysenteric bacilli: brief note by Mr. F. D'Herelle, presented by Mr. Roux. 1917. *Res Microbiol* 2007;**158**:553–4.
82. **d'Herelle F**. Bacteriophage as a treatment in acute medical and surgical infections. *Bull N Y Acad Med* 1931;**7**:329–48.
83. **Pirnay JP**, De Vos D, Verbeke G, *et al*. The phage therapy paradigm: pre-a-porter or sur-mesure? *Pharm Res* 2011;**28**:934–7.
84. **Kutateladze M**, Adamia R. Phage therapy experience at the Eliava Institute. *Med Mal Infect* 2008;**38**:426–30.
85. **Sulakvelidze A**, Alavidze Z, Morris JG Jr. Bacteriophage therapy. *Antimicrob Agents Chemother* 2001;**45**:649–59.
86. **Bruttin A**, Brussow H. Human volunteers receiving *Escherichia coli* phage T4 orally: a safety test of phage therapy. *Antimicrob Agents Chemother* 2005;**49**:2874–8.
87. **Gill JJ**, Hyman P. Phage choice, isolation, and preparation for phage therapy. *Curr Pharm Biotechnol* 2010;**11**:2–14.
88. **Raes J**, Letunic I, Yamada T, *et al*. Toward molecular trait-based ecology through integration of biogeochemical, geographical and metagenomic data. *Mol Syst Biol* 2011;**7**:473.
89. **Wang Y**, Antonopoulos DA, Zhu X, *et al*. Laser capture microdissection and metagenomic analysis of intact mucosa-associated microbial communities of human colon. *Appl Microbiol Biotechnol* 2010;**88**:1333–42.
90. **Mondot S**, Barreau F, Al Nabhani Z, *et al*. Altered gut microbiota composition in immune-impaired Nod2^{-/-} mice. *Gut*. Published Online First: 25 August 2011. doi:10.1136/gutjnl-2011-300478
91. **Petnicki-Ocwieja T**, Hrcir T, Liu YJ, *et al*. Nod2 is required for the regulation of commensal microbiota in the intestine. *Proc Natl Acad Sci U S A* 2009;**106**:15813–18.
92. **Rehman A**, Sina C, Gavrilova O, *et al*. Nod2 is essential for temporal development of intestinal microbial communities. *Gut* 2011;**60**:1354–62.
93. **Joossens M**, Huys G, Cnockaert M, *et al*. Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. *Gut* 2011;**60**:631–7.
94. **Van de Merwe JP**, Stegeman JH, Hazenberg MP. The resident faecal flora is determined by genetic characteristics of the host. Implications for Crohn's disease? *Antonie Van Leeuwenhoek* 1983;**49**:119–24.



A metagenomic insight into our gut's microbiome

Patricia Lepage, Marion C Leclerc, Marie Joossens, et al.

Gut 2013 62: 146-158 originally published online April 23, 2012
doi: 10.1136/gutjnl-2011-301805

Updated information and services can be found at:
<http://gut.bmj.com/content/62/1/146.full.html>

These include:

References

This article cites 92 articles, 37 of which can be accessed free at:
<http://gut.bmj.com/content/62/1/146.full.html#ref-list-1>

Article cited in:
<http://gut.bmj.com/content/62/1/146.full.html#related-urls>

Email alerting service

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Topic Collections

Articles on similar topics can be found in the following collections

[Gut Education](#) (56 articles)
[GUT Recent advances in basic science](#) (65 articles)

Notes

To request permissions go to:
<http://group.bmj.com/group/rights-licensing/permissions>

To order reprints go to:
<http://journals.bmj.com/cgi/reprintform>

To subscribe to BMJ go to:
<http://group.bmj.com/subscribe/>