



**Conference on ‘Diet, gut microbiology and human health’**  
**Symposium 1: The gut microbiome: forming the microbiome and modern approaches to its study**

**Functional metagenomics to decipher food–microbe–host crosstalk**

Pierre Larraufie<sup>1,2</sup>, Tomas de Wouters<sup>1,2</sup>, Gabrielle Potocki-Veronese<sup>3,4</sup>, Hervé M. Blottière<sup>1,2,5</sup>  
and Joël Doré<sup>1,2,5\*</sup>

<sup>1</sup>INRA, UMR 1319 Micalis, 78352 Jouy en Josas, France

<sup>2</sup>AgroParisTech, UMR Micalis, 78352 Jouy en Josas, France

<sup>3</sup>Université de Toulouse; INSA, UPS, INP; LISBP, 135 Avenue de Rangueil, F-31077 Toulouse, France

<sup>4</sup>UMR5504, UMR792 Ingénierie des Systèmes Biologiques et des Procédés, CNRS, INRA, F-31400 Toulouse, France

<sup>5</sup>INRA, US 1367 MetaGenoPolis, 78352 Jouy en Josas, France

The recent developments of metagenomics permit an extremely high-resolution molecular scan of the intestinal microbiota giving new insights and opening perspectives for clinical applications. Beyond the unprecedented vision of the intestinal microbiota given by large-scale quantitative metagenomics studies, such as the EU MetaHIT project, functional metagenomics tools allow the exploration of fine interactions between food constituents, microbiota and host, leading to the identification of signals and intimate mechanisms of crosstalk, especially between bacteria and human cells. Cloning of large genome fragments, either from complex intestinal communities or from selected bacteria, allows the screening of these biological resources for bioactivity towards complex plant polymers or functional food such as prebiotics. This permitted identification of novel carbohydrate-active enzyme families involved in dietary fibre and host glycan breakdown, and highlighted unsuspected bacterial players at the top of the intestinal microbial food chain. Similarly, exposure of fractions from genomic and metagenomic clones onto human cells engineered with reporter systems to track modulation of immune response, cell proliferation or cell metabolism has allowed the identification of bioactive clones modulating key cell signalling pathways or the induction of specific genes. This opens the possibility to decipher mechanisms by which commensal bacteria or candidate probiotics can modulate the activity of cells in the intestinal epithelium or even in distal organs such as the liver, adipose tissue or the brain. Hence, in spite of our inability to culture many of the dominant microbes of the human intestine, functional metagenomics open a new window for the exploration of food–microbe–host crosstalk.

**Metagenomics: Gut microbiota: Glycoside hydrolases: Bacterial functions**

Immune-mediated, chronic diseases of modern societies, steadily increasing worldwide since the 1950s, are most often associated with an alteration of the intestinal microbiota. The relevance of this observation is often challenged by the question of causality; i.e. the chicken or egg question. It nevertheless appears quite plausible that whether dysbiosis comes as a primary event or as a consequence of disease onset, in most pathophysiological conditions

alteration of intestinal ecology will promote stress signals for the immune system while aggravation of the inflammatory tone will deliver oxidative stress signals to the gut microbiota, such that a vicious circle is most likely to establish<sup>(1)</sup>. Preventing or tackling this context will hence require combining measures of immune and microbiota modulation. Both necessitate a finer understanding of food–microbe and microbe–host crosstalk.

\*Corresponding author: J. Doré, fax +33 1 3465 2070, email joel.dore@jouy.inra.fr



In a context where a large fraction of dominant gut microbes has remained inaccessible by classical culture methods<sup>(2)</sup>, functional metagenomics is providing the best approach to decipher food–microbe–host crosstalk. Shotgun sequencing of genetic information extracted from the intestinal environment allowed us to broadly revisit our view of the ecosystem in terms of phylogenetic structure and yielded massive datasets that may inform on functional potentials. At the same time, it has become possible to extract and clone large metagenomic fragments and directly explore their functionality with respect to food–microbe and microbe–host interfaces. Heterologous gene expression hence became a key for the culture-independent functional exploration of complex ecosystems relying on the extraction of environmental genomic DNA and its cloning in well-known cultivable hosts via fosmids, cosmids or bacterial artificial chromosomes. This exploration of functionalities of uncultivable bacteria within the metagenome of complex ecosystems is commonly referred to as activity-based functional metagenomics, as opposed to sequence-based phylogenetic or metagenomic characterisation of microbial ecosystems<sup>(3)</sup>.

#### Functional metagenomics and the food–microbe interface

One of the first applications of functional metagenomics gave access to formerly unknown fibrolytic enzymes from the microbiota of a thermophilic digester. Using *Escherichia coli* (a Gram-negative bacterium) as a host, even genes from anaerobic Gram-positive bacteria could be expressed<sup>(4)</sup>. Still using an *E. coli* host and a bacterial artificial chromosome library, Walter *et al.* screened a mouse intestinal metagenome for  $\beta$ -glucanase activity<sup>(5)</sup>. Out of a total of 5760 clones (representing 320 Mb genomic DNA, 55 kb per clone) they identified three clones encoding enzymes of interest. Similarly, Ferrer *et al.* screened a metagenome derived from rumen contents<sup>(6)</sup>. Within a metagenomic library of 14 000 clones (representing 77 Mb genomic DNA, up to 8 kb per clone) they identified and characterised twenty-two clones with distinct hydrolytic activities. In these two studies, the coverage of the actual metagenome was extremely limited owing to the size of the library. They nonetheless highlighted the ability of the functional approach relying on large-insert libraries (i.e. 10–50 kb) to yield access to full operons and operational gene clusters.

Jones *et al.* screened about 90 000 fosmid clones derived from a human faecal sample (representing 3.6 Gb large inserts metagenomic DNA) for bile salt hydrolase activity<sup>(7)</sup>. They observed these functions in all major gut bacterial divisions. In the same way, Tasse *et al.* screened 156 000 clones from a human faecal sample (representing 5.5 Gb metagenomic DNA) for hydrolytic activities towards different polysaccharides<sup>(8)</sup>. This exhaustive analysis allowed the identification of seventy-three novel carbohydrate-active enzymes, belonging to thirty-five known and nine totally novel glycoside-hydrolase families involved in the catabolism of complex

glycans in the human intestinal tract. Yet unknown mechanisms and three-dimensional structures have been pointed out for twelve families and fifteen families, respectively. This allowed the first discovery of novel functions of carbohydrate breakdown by using metagenomics<sup>(9)</sup>. Following a similar strategy, prebiotic-catabolising metagenomic clones were selected both from faecal and ileal samples, highlighting the potential of still unknown gut bacteria to metabolise functional foods<sup>(10)</sup>.

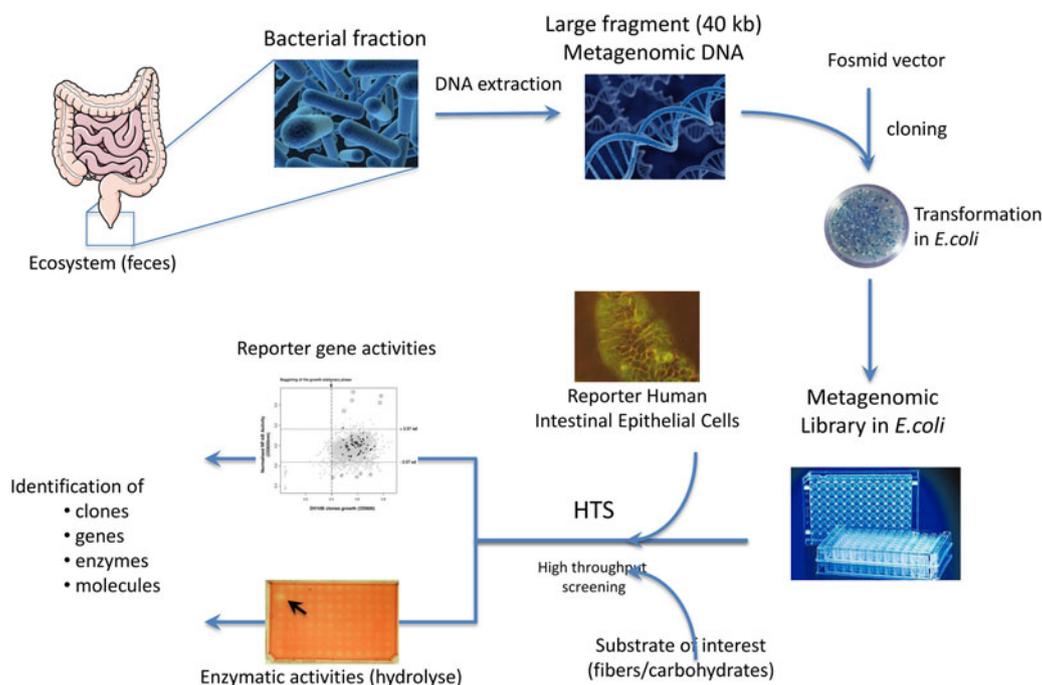
#### Functional metagenomics and the microbe–cell interface

Another important function of microbiota is its interaction or crosstalk with its host through intestinal epithelial cells. The use of reporter cell lines that allow quantification of specific functions such as cell proliferation, transcription factor activity or modulation of gene expression makes the screening of intestinal microbiota for these functions possible in a high-throughput format.

Interaction between microbes and eukaryotic cells can be explored using cultured commensal bacteria. Kaci *et al.* used a reporter system for a key transcription factor involved in the inflammatory response, NF  $\kappa$ B, to study the anti-inflammatory effects of different strains of *Streptococcus salivarius*<sup>(11)</sup>. Using human epithelial cells that were stably transfected with a reporter gene (in this case *Luciferase*) expressed under the control of a NF  $\kappa$ B binding domain, strains were tested for their capacity to down-regulate TNF  $\alpha$ -induced NF  $\kappa$ B activation simply by measuring reporter gene activity. Strains can be differentiated for their anti-inflammatory capabilities and characterisation of cellular mechanisms involved can be further implemented. This strategy has been used to test different bacterial strains of the species *Lactobacillus delbrueckii*<sup>(12)</sup>, questioning their anti-inflammatory properties as well as some pathogen activities<sup>(13)</sup>.

However, since culturable bacteria only represent a minority of total bacterial species of the intestinal ecosystem, metagenomic libraries may be used to find functions of unknown genes towards epithelial reporter cell lines in order to overpass this limitation, in analogy to their use to study enzymatic activities (Fig. 1). A first approach was to screen for cellular growth using crystal violet. In the present study, Gloux *et al.* screened over 20 000 metagenomic clones for modulation of proliferation of two human cell lines, HT-29 (modelling intestinal epithelial cells) and CV1 (a kidney fibroblast)<sup>(14)</sup>. This approach resulted in the identification of over fifty clones modulating cell proliferation. Interestingly, clones came from all major phyla represented in the two libraries tested in similar proportion, and a majority of genes belonged to unculturable bacteria, showing the possibilities offered by this technique<sup>(14)</sup>.

We subsequently used the NF  $\kappa$ B reporter system in two different human intestinal cell lines for functional metagenomic studies, and revealed new clones involved in modulation of this immunomodulatory pathway<sup>(15)</sup>. Screening of a library of mutants resulting from the random insertion of transposons in clones of interest was



**Fig. 1.** (Colour online) Schematic overview of the functional metagenomic approach developed to search for bacterial genes able to hydrolyse substrates of interest or to activate cell-signalling pathways in reporter human intestinal epithelial cells. Briefly, after isolation of faecal microbiota and DNA extraction, metagenomic DNA of about 40 kb are separated by pulse-field gel electrophoresis before cloning in a fosmid vector and infection of *E. coli* carrier bacteria to obtain metagenomic libraries. High-throughput screening (HTS) is then performed using robotic liquid-handling equipment to identify clones presenting the activity of interest, i.e. degradation of selected carbohydrates or modulation of reporter gene expression in human cells.

used to identify which genes of the clones were responsible for the activity. Use of reporter system with easily measurable activity such as luminescence, absorbance or fluorescence enables automation of screening, making possible use of larger libraries, representing a better coverage of the gut metagenome. Moreover, to decipher new microbe–host interactions, other screening models and advanced methods have been developed to study transcription factor activity such as activator protein 1 or PPAR $\gamma$  or gene expression related to the production of transforming growth factor  $\beta$ , indoleamine 2,3-dioxygenase, thymic stromal lymphopoietin and others<sup>(16)</sup>. Recent advances in eukaryotic genome engineering with development of TALEN or Crispr/Cas9 technologies may be used for the development of new reporter systems by tagging endogenous genes. The development of these new tools will allow the discovery of new functionally bioactive clones, for which mechanistic studies will further expand our knowledge of interactions between the microbiota and its host.

However, one major drawback of functional metagenomics is that genes encoded in clones may not be expressed functionally, as they may require specific maturation or secretion systems. Moreover, bacterial genes are often expressed in functional clusters. This is why most screenings published so far have used large inserts, from 10 to 50 kb, to allow expression of potentially full operons required for the expression of a gene and

observation of its associated activity. Still, whole organism and not only adjacent genes may be important for their functional expression. Until now, *E. coli* has been the only model of bacteria used for gut microbiota screening, but other host models should also be developed. Even though functional genomics using *E. coli* as the host strain made it possible to discover bioactive genes from Gram-positive bacteria, only about 40 % of heterologous genes are functionally expressed in this strain<sup>(17)</sup>. Different efforts to use a Gram-positive host strain have been made, as most probiotics (which may by definition confer beneficial effects on the host) are Gram-positive, and this kind of bacteria has specific secretion and cell-wall anchoring systems. Dobrijevic *et al.* recently developed a strategy to use a *Bacillus subtilis* strain with low basal activity as a vector of expression<sup>(18)</sup>. However, the use of this model for high throughput screening has not been published yet.

As a conclusion, metagenomics allows a good and exhaustive characterisation of the human intestinal ecosystem going from who is there to what do they do, towards an ecological understanding of the gut<sup>(1)</sup>. The functional metagenomic approach is an interesting complement to full sequencing. Deeper characterisation of the functions of the microbiome in a healthy state as well as in pathological conditions in which microbiota dysbiosis is observed remains to be investigated further.

### Financial Support

G. P. V. and J. D. were members of ANR consortium AlimIntest (ANR-06-PNRA-0024). T. D. W., J. D. and H. M. B. were members of the European Marie-Curie Initial Training Network Cross-Talk (grant agreement #215553). T. D. W. was supported by the ANR blanc MicroObes (ANR-07-GMGE-002.1-01). This work was supported by the ANR blanc FunMetaGen (ANR-11-BSV6-0013); by the European Union's Seventh Framework Programme, Metagenomics of the Human Intestinal Tract; HEALTH-F4-2007-201052 and Metagenomics in Cardiometabolic Diseases; HEALTH-F4-2012-305312.

### Conflicts of Interest

None.

### Authorship

All authors contributed to the writing and editing processes.

### References

1. Blottière HM, de Vos WM, Ehrlich SD *et al.* (2013) Human intestinal metagenomics: state of the art and future. *Curr Opin Microbiol* **16**, 232–239.
2. Suau A, Bonnet R, Sutren M *et al.* (1999) Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. *Appl Environ Microbiol* **65**, 4799–4807.
3. Handelsman J (2004) Metagenomics: application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* **68**, 669–685.
4. Healy FG, Ray RM, Aldrich HC *et al.* (1995) Direct isolation of functional genes encoding cellulases from the microbial consortia in a thermophilic, anaerobic digester maintained on lignocellulose. *Appl Microbiol Biotechnol* **43**, 667–674.
5. Walter J, Mangold M & Tannock GW (2005) Construction, analysis, and beta-glucanase screening of a bacterial artificial chromosome library from the large-bowel microbiota of mice. *Appl Environ Microbiol* **71**, 2347–2354.
6. Ferrer M, Golyshina OV, Chernikova TN *et al.* (2005) Novel hydrolase diversity retrieved from a metagenome library of bovine rumen microflora. *Environ Microbiol* **7**, 1996–2010.
7. Jones BV, Begley M, Hill C *et al.* (2008). Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome. *Proc Natl Acad Sci USA* **105**, 13580–13585.
8. Tasse L, Bercovici J, Pizzut-Serin S *et al.* (2010) Functional metagenomics to mine the human gut microbiome for dietary fiber catabolic enzymes. *Genome Res* **20**, 1605–1612.
9. Ladevèze S, Tarquis L, Cecchini DA *et al.* (2013) Role of glycoside-phosphorylases in mannose foraging by human gut bacteria. *J Biol Chem* **288**, 32370–32383.
10. Cecchini DA, Laville E, Laguerre S *et al.* (2013) Functional metagenomics reveals novel pathways of prebiotic breakdown by human gut bacteria. *PLoS ONE* **8**, e72766.
11. Kaci G, Lakhdari O, Doré J *et al.* (2011) Inhibition of the NF-kappaB pathway in human intestinal epithelial cells by commensal *Streptococcus salivarius*. *Appl Environ Microbiol* **77**, 4681–4684.
12. Santos Rocha C, Lakhdari O, Blottière HM *et al.* (2012) Anti-inflammatory properties of dairy lactobacilli. *Inflamm Bowel Dis* **18**, 657–666.
13. Madi A, Lakhdari O, Blottière HM *et al.* (2010) The clinical *Pseudomonas fluorescens* MFN1032 strain exerts a cytotoxic effect on epithelial intestinal cells and induces Interleukin-8 via the AP-1 signaling pathway. *BMC Microbiol* **10**, 215.
14. Gloux K, Leclerc M, Iliozer H *et al.* (2007) Development of high-throughput phenotyping of metagenomic clones from the human gut microbiome for modulation of eukaryotic cell growth. *Appl Environ Microbiol* **73**, 3734–3737.
15. Lakhdari O, Cultrone A, Tap J *et al.* (2010) Functional metagenomics: a high throughput screening method to decipher microbiota-driven NF- $\kappa$ B modulation in the human gut. *PLoS ONE* **5**, e13092.
16. de Wouters T, Ledue F, Nepelska M *et al.* (2014) A robust and adaptable high throughput screening method to study host–microbiota interactions in the human intestine. *PLoS ONE* **9**, e105598.
17. Gabor EM, Alkema WBL & Janssen DB (2004) Quantifying the accessibility of the metagenome by random expression cloning techniques. *Environ Microbiol* **6**, 879–886.
18. Dobrijevic D, Di Liberto G, Tanaka K *et al.* (2013) High-throughput system for the presentation of secreted and surface-exposed proteins from Gram-positive bacteria in functional metagenomics studies. *PLoS ONE* **8**, e65956.