Functional Metagenomic Investigations of the Human Intestinal Microbiota

Moore, Aimee M.; Munck, Christian; Sommer, Morten Otto Alexander; Dantas, Gautam

Published in:
Frontiers in Microbiology

Link to article, DOI:
10.3389/fmicb.2011.00188

Publication date:
2011

Document Version
Publisher’s PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Functional metagenomic investigations of the human intestinal microbiota

Aimee M. Moore1,2†, Christian Munck3†, Morten O. A. Sommer3* and Gautam Dantas1,4*

1 Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, MO, USA
2 Department of Pediatrics, Washington University School of Medicine, St. Louis, MO, USA
3 Department of Systems Biology, Technical University of Denmark, Lyngby, Denmark
4 Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO, USA

Keywords: functional metagenomics, human intestinal microbiota, antibiotic resistome

The human intestinal microbiota encode multiple critical functions impacting human health, including metabolism of dietary substrate, prevention of pathogen invasion, immune system modulation, and provision of a reservoir of antibiotic resistance genes accessible to pathogens. The complexity of this microbial community, its recalcitrance to standard cultivation, and the immense diversity of its encoded genes has necessitated the development of novel molecular, microbiological, and genomic tools. Functional metagenomics is one such culture-independent technique, used for decades to study environmental microorganisms, but relatively recently applied to the study of the human commensal microbiota. Metagenomic functional screens characterize the functional capacity of a microbial community, independent of identity to known genes, by subjecting the metagenome to functional assays in a genetically tractable host. Here we highlight recent work applying this technique to study the functional diversity of the intestinal microbiota, and discuss how an approach combining high-throughput sequencing, cultivation, and metagenomic functional screens can improve our understanding of interactions between this complex community and its human host.
and thereby provide a rich source for mining novel therapeutic approaches.

To understand microbial community effects on human health, both the phylogenetic profile of human microbial communities and the functional capacity of their members must be characterized. Much progress has been made toward these ends using direct bacterial culture, 16S sequencing, shotgun metagenomic sequencing, PCR probing for specific genes, and chemical profiling of microbial metabolites. These approaches have yielded incredible insights ranging from shifts in prevalent bacterial phylotypes and altered metabolic profiles in human subjects with inflammatory bowel disease, variations in the composition of the intestinal microbiota with human diet and functional differences in the gut microbiota related to host body habitus, developmental changes in the composition of the gastrointestinal microbiota during infancy and childhood, and the genetic epidemiology of antibiotic resistance in the intestinal microbiota. (Rimbara et al., 2005; Qin et al., 2006; Turnbaugh et al., 2006; Bezeab et al., 2009; Jansson et al., 2009; Paliy et al., 2009; Gillevert et al., 2010; Kang et al., 2010; Koenig et al., 2011; Rigsbee et al., 2011). In this perspective, we will focus on the emerging application of functional metagenomic screens, a technique developed for investigating unculturable environmental microbes that neatly complements the aforementioned techniques currently used to characterize the human microbiota.

Direct culture, historically the sine qua non of microbiology, readily provides information on the functional characteristics of the species being investigated. The majority of gastrointestinal microbiota, however, are obligate anaerobes recalcitrant to culture. Traditional estimates are that only 15–20% of the gastrointestinal microbiota are culturable, precluding direct characterization of the majority of bacterial species (Langendijk et al., 1995; Eckburg et al., 2005; Gill et al., 2006). A recent report by Goodman et al. (2011) showed, using high-throughput 16S sequencing in combination with extensive anaerobic culturing, that up to 56% of gastrointestinal microbial species are culturable. Although this represents a dramatic improvement over standard culturing techniques, there remains a significant proportion of unculturable organisms that must be characterized using complementary techniques. Different approaches have been employed to overcome this problem ranging from simple PCR-based screens to large metagenomic sequencing analyses and functional metagenomic screens. Together, these methods have expanded our knowledge about the fraction of the GI tract microbiota that cannot be characterized by culture-based approaches.

**FUNCTIONAL METAGENOMICS: AN EMERGING TECHNIQUE FOR CHARACTERIZING UNCULTURABLE ORGANISMS**

Functional metagenomic screens, originally proposed as a method to characterize the unculturable fraction of soil microbiota (Handelsman et al., 1998; Rondon et al., 2000) and successfully used for years to characterize the functional diversity of microbes in a variety of environments (Wärnecke et al., 2007; Allen et al., 2009b; Berlemont et al., 2009; Torres-Cortes et al., 2011), has relatively recently been adapted to characterize the functions of human microbial communities, representing an interesting cross-pollination between environmental microbiology and biomedical science. The functional metagenomic screening method, based on clone libraries containing genomic DNA from a microbial community, does not require direct culture of fastidious organisms. Instead, clone libraries are constructed by first extracting and shearing DNA from a sample of a microbial community, then cloning the fragmented DNA into a relevant vector, and subsequently transforming this vector into a suitable host strain (Figure 1). Once a library is constructed, it can be functionally screened by cultivation on selective media or by employing a reporter system. Using this approach, it is possible to identify harboring DNA fragments that encode antibiotic resistance genes are selected by subjecting the library of clones to specific antibiotics at concentrations which inhibit the growth of the untransformed indicator strain. Selected DNA fragments can then be sequenced to identify the specific resistance genes.

![Functional Metagenomics Workflow](image-url)
genes encoding a variety of functions such as antibiotic resistance, metabolism of complex compounds, and modulation of eukaryotic cells. Subsequent sequencing and in silico analysis of the DNA inserts from isolated clones provides information about the source of the genes and the putative mechanisms of action of their products.

INTERACTIONS WITHIN MEMBERS OF THE INTESTINAL MICROBIOTA: ANTIBIOTIC RESISTANCE

One area of early success for functional metagenomic screens is in the discovery of new antibiotic resistance genes in the human gastrointestinal microbiota. Multidrug resistant bacteria are increasingly prevalent in both hospitals and the community, and pose a growing threat to human health (Boucher et al., 2009; Högb erg et al., 2010). Infections with antibiotic resistant organisms are associated with increased mortality and cost of treatment (Maragakis et al., 2008), and novel antibiotic discovery has not kept pace with the emergence of microbial resistance to existing agents (Högb erg et al., 2010). In order to develop a rational approach to curtail the emergence of antibiotic resistance in human pathogens, a deeper understanding of the flow of resistance genes within microbial communities is required. Pathogenic organisms present in the environment may acquire resistance genes from soil or water microbes, while commensal gastrointestinal organisms that are continuously exposed to the outside environment via host ingestion of food, may also come in contact with pathogenic bacteria during the course of an infection. Although great strides have been made in recent years documenting genetic resistance reservoirs and patterns of gene flow within and between environmental and human commensal microbiota, fully characterizing the diversity and mobility of the environmental resistome will be crucial to control the emergence of ever more resistant organisms (Aminov and Mackie, 2007; Martinez, 2008; Aminov, 2009; Allen et al., 2010).

Multiple studies demonstrate the efficacy of simple PCR screens in the detection and quantification of antibiotic resistance genes present in the gastrointestinal microbiota. PCR assays have been used to detect the presence of known tetracycline resistance genes (tet) in fecal samples from antibiotic-naïve infants (Gueimonde et al., 2006). Similarly, qPCR has been used to quantify the levels of tet and erm genes, which confer resistance to tetracycline and macrolide, lincosamide, and streptogramin B antibiotics respectively, in animal and human waste water (Smith et al., 2004; Auerbach et al., 2007; Chen et al., 2010). The extraordinary specificity of PCR-based studies is also an important limitation of the technique: because PCR can only be used to interrogate a sample for known genes, it is an ineffective method for identifying novel resistance genes.

Functional metagenomic screens obviate this problem by identifying genes by their function in an expression vector rather than by a specific sequence used for PCR probing. Using this approach, novel antibiotic resistance genes have been identified in different environments including oral microbiota, soil microbiota, and moth gut flora (Diaz-Torres et al., 2003; Riesenfeld et al., 2004; Allen et al., 2009a). Sommer et al. (2009) demonstrated the power of metagenomic functional screens to identify novel antibiotic resistance genes in fecal samples from two healthy adults. Metagenomic libraries with a total size of 9.3 Gb (gigabases) and an average insert size of 1.8 kb (kilobases) were screened for resistance against 13 different antibiotics, revealing 95 unique inserts representing a variety of known resistance genes as well as 10 novel beta-lactamase gene families (Sommer et al., 2009). Genes identified using metagenomic functional screens were, on average, 61% identical to known resistance genes from pathogenic organisms, while genes identified via aerobic culturing of isolates from the same individuals had greater than 90% sequence identity to previously described resistance genes. One of the novel resistance genes identified with the functional metagenomic screen had 100% sequence identity to a protein of unknown function, demonstrating the power of metagenomic functional screens to identify novel resistance genes even in fully sequenced and apparently well-annotated organisms. Antibiotic resistance with high sequence identity to known genes were more likely than novel genes to be flanked by mobile genetic elements such as transposases, possibly indicating that the novel genes represent a potential resistance reservoir that has not yet become widely disseminated. Recent work by Goodman et al. (2011) demonstrated that interindividual differences in gastrointestinal antibiotic resistance genes can be detected by subjecting both uncultured fecal samples as well as pools of phylogenetically representative fecal culture collections to metagenomic functional screens. Notably, the presence or absence of specific resistance genes (e.g., those encoding amikacin resistance) in uncultured samples, as determined by functional metagenomics, correlated with the fraction of cultured isolates phenotypically resistant to those compounds, and the presence of the exact genes identified by functional metagenomic screens was reconfirmed by PCR assay in phenotypically resistant cultured strains. These authors also found that the nearest genome-sequenced phylogenetic neighbors of the resistant strains isolated from the the gastrointestinal microbiota of the sampled individuals lacked similar resistance genes, further highlighting the diversity and individualized nature of antibiotic resistomes. Together, these studies indicate that the gastrointestinal microbiota are likely to harbor many more resistance genes that will continue to be revealed by further investigation.

Functional metagenomic screens have also been used to mine the resistance reservoir in the intestinal microbiota of farm animals. Livestock are frequently dosed with antibiotics to treat infections and promote growth, and mounting evidence indicates that these practices lead to increased antibiotic resistance not only in the microbiota of the treated animals but also in their human caregivers (Sorum et al., 2006). The scope of this problem is highlighted by the findings of Kazmierczak et al. (2009), who employed metagenomic functional screens to identify both known and novel tetracycline resistance genes in fecal samples from organically farmed pigs that had not been exposed to antibiotics. Most of these genes were associated with mobile genetic elements, possibly explaining their persistence in an environment without any obvious selection pressure. The clinical and epidemiologic relevance of resistance genes present in the intestinal microbiota must be further defined by examining secondary effects of these genes, such as fitness costs or benefits associated with their expression, as well as by demonstrating the potential for direct transfer of the resistance gene to pathogenic organisms.
FUNCTIONAL METAGENOMICS FOR UNDERSTANDING THE GENETIC DETERMINANTS OF METABOLIC FUNCTION IN THE GASTROINTESTINAL MICROBIOTA

As previously noted, specific variations in the composition of the gastrointestinal microbial community have been linked to important states of human health and disease. Recent advances in understanding the interactions between bacterial metabolites and the host cellular machinery have begun to illuminate the physiologic basis of microbial contributions to human pathology. Metabolites generated either directly or indirectly by saccharolytic species may provide a mechanistic explanation for the observed human health outcomes associated with different compositions of the enteric microbial community. Conjugated linoleic acids, which are generated by some Bifidobacterial species (Rosberg-Cody et al., 2004), modulate tumorigenesis in animal models (Kelley et al., 2007), and are being investigated for a role in modulating inflammation and risk for neoplasia in humans (Bhattacharya et al., 2006; Coakley et al., 2009). Short-chain fatty acids (SCFAs) are bacterial metabolites that have wide-ranging effects on human physiology. In animal models of prematurity, some SCFAs (acetic and butyric acid) directly injure colonic mucosa in a dose-dependent fashion in the most immature age groups (Lin et al., 2002), an effect that disappears with increasing postnatal age (Nafday et al., 2005). This suggests a possible role for bacterial metabolites in the complex pathogenesis of necrotizing enterocolitis, a necro-inflammatory disease commonly seen in preterm infants but non-existent in older age groups. Butyrate, a SCFA that is produced by fermentation of dietary fiber, has a variety of effects modulating inflammation and risk for neoplasia (Hamer et al., 2008). Butyrate is taken up by colonocytes via the MCT1 and SMCT1 transporters (downregulated in cancer cells), and is protective against colon cancer in animal models. It also inhibits histone deacetylase and inhibits TNF-αB activation, which may explain its role in modulating inflammation. Acetate and propionate, two other SCFAs, have opposing effects on cholesterol biosynthesis (Wong et al., 2006). Microbe-generated SCFAs also may contribute to host obesity via interaction with fasting-induced adipocyte factor (Fiaf), AMPK, and Gpr41, which modulate pathways regulating fatty acid storage in adipocytes, fatty acid oxidation, gastrointestinal motility, and nutrient absorption (Backhed et al., 2007; Samuel et al., 2008).

Functional metagenomic screens offer a powerful means for detecting the genetic determinants of microbial metabolism. Jones et al. (2008) employed a functional screen using a large-insert metagenomic library to identify bile salt hydrolases within the human gastrointestinal microbiota. End-sequencing of clones displaying bile salt hydrolase activity revealed a broad phylogenetic distribution of bile salt hydrolase enzymes suggesting that this metabolic capacity is a conserved trait among bacteria adapted to life in the human gastrointestinal tract. Since bile salts play important roles in the processing and uptake of dietary fats in the intestines, microbial catabolism of these compounds may affect the amount of energy extracted from the diet.

Catabolism of fibers indigestible by the human host, another significant activity of the human intestinal microbiota, has been investigated using successive rounds of functional screens to enrich the metagenomic library with carbohydrate-metabolizing enzymes followed by high-throughput sequencing to identify genetic determinants of carbohydrate metabolism within the human gastrointestinal microbiota (Tasse et al., 2010). They identified 73 carbohydrate-metabolizing enzymes from the enriched library, representing a fivefold increase in active genes identified compared to metagenomic sequencing without enrichment. This highlights the strong potential of serial functional screens combined with high-throughput sequencing to identify novel genes and yield increasingly comprehensive information on the metabolic potential of a given microbial community.

INTEGRATING FUNCTIONAL SCREENS WITH SHOT-GUN METAGENOMIC SEQUENCING ANALYSIS

The advent of convenient applications for metagenomic data analysis such as MG-RAST and MEGAN have simplified annotation and comparative analysis of functionally selected genes, which together with the declining cost of high-throughput sequencing, offer an efficient complement to functional screens (Huson et al., 2007; Meyer et al., 2008). Several studies have used this approach to connect functional genes with metabolic capacities and to identify pathways, such as metabolism of sugars, amino acids, and nucleotides, that are enriched in the gastrointestinal microbiota relative to representative genome-sequenced strains (Gill et al., 2006; Kurokawa et al., 2007; Turnbaugh et al., 2009; Arumugam et al., 2011). Moreover, by ranking functional gene clusters according to their frequencies, a minimal gut genome and a minimal gut metagenome have been described (Qin et al., 2010). The former reflects the minimal set of genes required by a single member of the gastrointestinal microbiota, while the latter indicates the minimal set of genes required to sustain the aggregate gastrointestinal microbiota. The minimal gut genome includes genes essential to all bacteria (e.g., replication, transcription, translation) as well as gut-specific genes such as those facilitating adhesion to epithelium. In contrast, the minimal gut metagenome includes genes necessary for metabolism of complex sugars, underscoring the importance of coupled metabolism in sustaining the GI tract microbiota. The importance of confirming gene function in vitro and in vivo to ensure reliable annotation is illustrated by Hess et al. (2011), who used metagenomic sequencing to identify >20,000 carbohydrate active genes from the cow rumen microbiota. From this gene set, they selected 90 in silico predicted carbohydrate-metabolizing genes, expressed them, subjected them to functional assays, and found that 51 genes were enzymatically active in vitro (Hess et al., 2011). These studies exemplify how metagenomic sequencing, automated annotation of large data sets, and functional screening comprise a powerful toolkit capable of characterizing functional networks in highly complex environments such as the GI tract microbiota.

FUNCTIONAL MAPPING OF INTERACTIONS BETWEEN HUMANS AND THEIR INTESTINAL MICROBIOTA

Functional metagenomic screens may also illuminate the genetic determinants of microbial interactions with host cells. The intestinal microbiota have long been known to modulate intestinal epithelia, for instance, by stimulating intestinal cell differentiation (Bry et al., 1996). In order to identify specific bacterial gene products that directly influence the fate of human cells, Gloux et al.
Functional metagenomic screens have also been designed to investigate the immune-modulatory capacity of the gastrointestinal microbiota. To identify clones modifying the host immune response, Lakhdari et al. (2010) constructed an NF-κB activated reporter system from a human colorectal carcinoma cell line. By screening metagenomic libraries of GI tract microbiota from patients with Crohn’s disease, in which NF-κB activity is frequently elevated (Ellis et al., 1998), they identified several clones either inducing or inhibiting NF-κB activity. Together, these studies demonstrate the potential for functional metagenomic screens to illuminate the genetic mechanisms for microbial community contribution to the development of the human immune system and the pathogenesis of atopic, autoimmune, and neoplastic disease, which may provide novel therapeutic targets for these conditions.

In addition to interacting with human cells, commensal bacteria can also use quorum-sensing to convey signals over distances and thereby coordinate community gene expression. Guan et al. (2007) used a metabolite regulated expression(METREX) screen based on a quorum-sensing inducible promoter fused to gfp to identify genes encoding a new class of quorum-sensing inducing molecules in moth gut microbiota, demonstrating the power of functional metagenomics for characterizing the determinants of community behavior in uncultured organisms.

**FUNCTIONAL METAGENOMICS FOR REFINING PRE- AND PRO-BIOTIC INTERVENTIONS**

Increased understanding of the effects of gastrointestinal microbiota on human health has generated interest in targeting these communities for therapeutic intervention (Cani and Delzenne, 2011). Short-chain carbohydrates that are indigestible by humans but are fermentable by some microbes have demonstrable efficacy in increasing the populations of Lactobacilli and Bifidobacteria in the human gastrointestinal tract (Wang and Gibson, 1993). Investigations of galactose oligosaccharides (GOS) and fructose oligosaccharides (FOS) as additives to infant formula have demonstrated increased Bifidobacterial populations, decreased stool pH, generation of a stool fatty acid profile more similar to that found in breastfed infants, and reduced populations of potential pathogens such as Clostridia spp., Bacteroides spp., and E. coli (Fanaro et al., 2005; Knol et al., 2005; Costalos et al., 2008; Magne et al., 2008; Rao et al., 2009). Prebiotic supplementation with oligosaccharides may promote blooms of beneficial bacteria more effectively than direct administration of pro-biotic organisms: a study directly comparing infant formula containing Bifidobacterium animalis with GOS/FOS-supplemented formula revealed a significantly greater proportion of Bifidobacterial species in the infants fed oligosaccharide-containing formula but no difference between the Bifidobacterial supplemented formula and control formula groups (Bakker-Zierikzee et al., 2005). Administration of prebiotics such as inulin and oligosaccharides in adult humans have shown some effect on hunger and satiety mechanisms (Whelan et al., 2006; Cani et al., 2009) but inconsistent results when applied to pathologies such as atopy and inflammatory bowel disease (Guarnier, 2005; Roberfroid et al., 2010). Functional metagenomics has the potential to refine current prebiotic therapies by more completely defining the genetic determinants of metabolism for given constituents of a microbial community, providing a rational basis for more precise design of prebiotic agents intended to promote blooming of a specific subset of organisms.

**TOWARD A COMPLETE FUNCTIONAL REPRESENTATION OF THE GASTROINTESTINAL MICROBIOTA**

Functional metagenomic screens have been successful in elucidating novel genes encoding microbial antibiotic resistance, metabolic machinery, and immune-modulatory elements. Despite their demonstrable utility, functional metagenomic screens have several important limitations. First, the DNA insert must be compatible with the host’s expression machinery and the gene product must be non-toxic and functional in the host (for an in-depth review, see Uchiyama and Miyazaki, 2009). Second, the host must be suited for the screen: when screening for antibiotic resistance genes, a host sensitive to the antibiotic of interest must be chosen. Third, the insert size may restrict the diversity of functions portrayed in a screen; a small insert library cannot reveal the function of genes organized in large operons such as many metabolic pathways or some efflux pumps associated with antibiotic resistance. Finally, the expression level of the insert can significantly affect the result of a functional screen. Using a high-copy plasmid as vector or a strong synthetic promoter can result in an overestimation of functionality. Conversely, overexpression of potentially lethal proteins may cause underestimation of functional genes, (e.g., cell lysis due to overexpression of efflux pumps). Despite these limitations, multiple studies demonstrate the potential for functional metagenomic screens to powerfully complement direct culture, 16S sequencing, shotgun metagenomic sequencing, and metabolomic analysis to offer new insight into the complex interactions between microbial communities and their human hosts. Used in concert, these techniques promise to expand our understanding of microbial community function, its impact on human health, and to provide novel targets for therapeutic development in the coming years.

**REFERENCES**


Aminov, R. I. (2009). The role of antibiotics and antibiotic resistance...


Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 01 April 2011; paper pending published: 31 May 2011; accepted: 23 August 2011; published online: 17 October 2011.


This article was submitted to Frontiers in Cellular and Infection Microbiology, a specialty of Frontiers in Microbiology. Copyright © 2011 Moore, Munck, Sommer and Dantas. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.