Network-Based Approach to Collective Microbial Activities in the Human Gut Ecosystem

PAN-JUN KIM
DEPARTMENT OF PHYSICS, KOREA ADVANCED INSTITUTE OF SCIENCE AND TECHNOLOGY

ABSTRACT

The role of our gut microbial ecosystem (gut microbiota) in health and disease is largely attributed to the collective metabolic activities of the inhabitant microbes. A global mechanistic framework of the microbial community structure, mediated through metabolite transport, would provide important insights into the collective dynamics governing host-associated microbial ecology. Recently, we have constructed a literature-curated network of human gut microbiota, with metabolic interactions between microbial species coupled with the host system. The network encompasses ~570 microbial species and three human cell types, interlinked by chemical compounds via >4,400 annotated small-molecule transport and macromolecule degradation events. We have demonstrated the utility of our network in the analysis of disease-related microbiota by computing metabolic influence in the gut ecosystems of type 2 diabetes patients. This analysis reveals the disease-specific infrastructure of the gut microbial ecosystem, core microbial groups with distinctively large metabolic influence, and possible metabolic compounds from disease-relevant community processes. Our network framework shows promise for investigating complex microbe-microbe and host-microbe chemical cross-talk, and for identifying disease-associated features.

INTRODUCTION

Microbial communities populate almost every part of the planet, ranging from hot springs to deep oceans, soil environments, and animal bodies. Central to community dynamics and evolution are the interactions of microbes with each other and with the environment. Among various microbial habitats, our intestine is the site of an extraordinarily complex and dynamic symbiosis. The adult human intestine harbors as many microbial cells as human cells [1]. Our resident gut microbial community, or gut microbiota (microbiome), provides us with a variety of biochemical capabilities not encoded in our genome, and exerts an important influence on many aspects of human physiology involved in health and disease. The most densely colonized area is the large intestine, wherein microbes survive and grow by consuming diet-derived and host-derived chemical compounds as well as metabolic byproducts excreted by other microbes. Undigested dietary macromolecules and host-derived substrates are broken down by microbial species, and then the solubilized molecules become available to other members of the community as public goods for uptake. Additionally, the inherent microbial activities of importing metabolic resources and exporting metabolic byproducts give rise to competition for metabolic resources and cooperative relations (such as cross-feeding of metabolic byproducts) among resident microorganisms in our gut environment. Furthermore, microbial metabolic waste products such as short-chain fatty acids (SCFAs) have active roles in normal host physiology, as energy sources for colonocytes, regulators of gene expression and cell differentiation, and anti-inflammatory agents. On the other hand, some metabolic byproducts exported by microbes can be highly toxic to host tissues and impair their function, promoting the onset and progression of disease [2]. These intricate microbe-microbe and microbe-host interconnections result in the formation of a complex ecological network in the human gut environment. The emergent community structures make evident that, notwithstanding the importance of individual microbial species, the microbiota

dynamics and consequent impact on the host can be largely attributed to the collective activities of numerous microbial species and metabolic compounds, thoroughly interlinked by network relationships. This realization calls for an integrative network-based approach for a system-level understanding of the human gut microbiota [3].

Advances in sequencing technologies and metagenomics have revealed significant associations between our gut microbiota and a number of human phenotypes and diseases, such as obesity, inflammatory bowel disease, colorectal cancer, and type 2 diabetes [2]. Type 2 diabetes (T2D), the most prevalent endocrine disease, represents one of the global health challenges of this century. A total of 387 million people in the world, i.e., 9% of the adult (>18 years old) population, have diabetes, of which 90% corresponds to T2D [4]. Among several risk factors for T2D, socio-demographic and environmental factors seem to have greater influence than genetic factors. Previous studies indicate that, as a prominent environmental factor, alternation in the gut microbiota might contribute to the pathology of this disease [5].

Recently, we have constructed a global metabolic interaction network of the human gut microbiota, NJS16 [3]. NJS16 is a comprehensive literature-curated microbiota interaction map that accounts for 567 bacterial and archaeal species in the large intestine, 244 metabolic compounds, 4,483 transport or degradation events of metabolic compounds, and three important human cell types metabolically interacting with microbes in the large intestine. The network architecture stands on mechanistic information that reflects the metabolic potential of the gut microbiota supported by current biological or experimental knowledge. We have demonstrated the utility of our network for the community analysis of the gut ecosystem involved in disease, specifically, T2D [3].

RESEARCH OUTCOMES

Global landscape of the human gut microbiota organized through metabolite transport

We aimed to construct a community-level network composed of microbial species populating the human gut. To construct our network, we started by applying a phylogenetic analysis tool on previously published, shotgun metagenomic sequencing data from fecal samples of Chinese individuals [5]. Next, we extensively searched the published literature for all annotated, mainly experimental, information that described the small-molecule metabolites (e.g., monosaccharides, disaccharides, SCFAs, vitamins, and gases), which are transported into, and/or out of, the microbial species identified in the microbiome samples. To complement our list of annotated microbe-metabolite associations, we included macromolecule degradation reactions, i.e., microbes and the macromolecules (e.g., cellulose, hemicellulose, inulin, starch, and mucin) that are known to degrade, as well as the resulting degradation products (e.g., D-glucose and cellobiose from cellulose; N-acetylglucosamine, N-acetylneuraminate, L-fucose, and sulfate from mucin). Although tissue cells of the human host are not physically a part of the gut microbiota, we viewed them as a functional extension of the bacterial and archaeal community residing in the colon, because host cells either can directly affect, or can be affected by, microbial metabolism. The specific

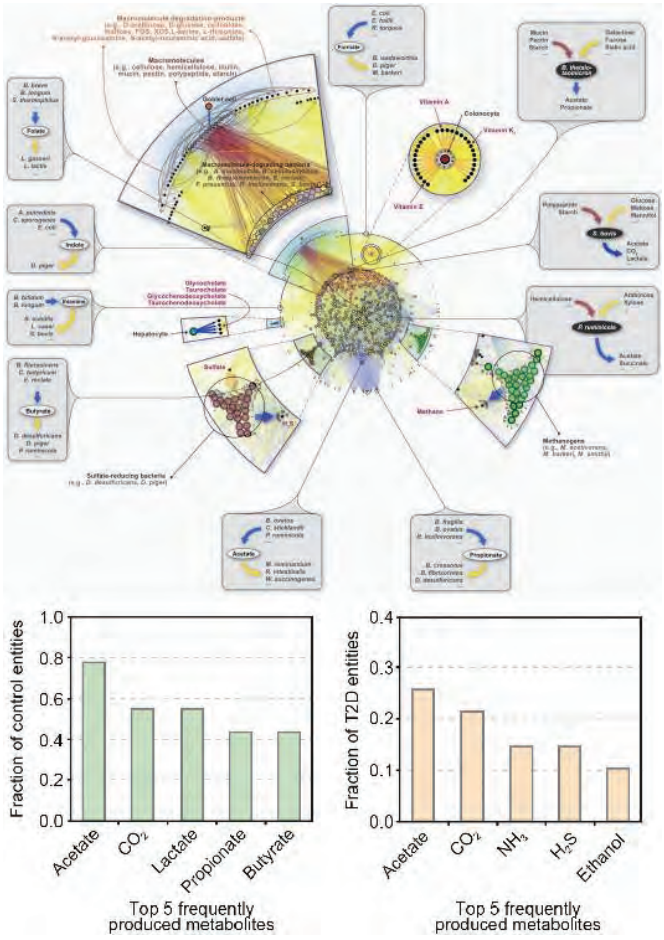

Fig. 1: (upper panel) Global metabolic interaction network of the human gut microbiota that we have constructed. Nodes represent microbial species, human cell types, and metabolic compounds, and links represent metabolite import, export, and degradation events of microbial species and human cell types. (bottom panels) Frequently produced metabolites from our gut microbiota network. Results from a nondiabetic control group (left) or T2D patient group (right).

host cells that we considered were the colonocyte, the goblet cell (for mucin secretion), and the hepatocyte (for glycine- or taurine-conjugated bile acid export). Lastly, we linked all microbes and host cells to their associated metabolites and macromolecules into a comprehensive reference map of human gut microbiota and chemical compound relationships. More specifically, a community member and a chemical compound were then connected by a (directed) link if the organism could import and/or export the metabolite, or degrade the macromolecule.

As a result, we constructed NJS16, a literature-curated community-level network of the human gut microbiota organized through metabolite transport (Fig. 1, upper panel) [3]. Our network is a collection of 4,483 transport or degradation reactions (from about 400 research articles, reviews, and textbooks) between 244 metabolic compounds (229 small molecules and 15 macromolecules) and 570 microbial species and human cell types (511 bacteria, 56 archaea, and 3 host cells). Clearly, metabolite-driven cooperative relationships are pervasively seen throughout the entire network (Fig. 1, upper panel).

Using NJS16, we can now explore the organizing characteristics of global microbial metabolic activities in the human gut environment.

Disease-associated microbial metabolic influences

Next, we demonstrated the potential of our global metabolite transport network of the gut microbiota for elucidating context-specific, community-level features. We developed a mathematical framework and applied it on the aforementioned Chinese microbiome samples, which were obtained from T2D patients and non-diabetic controls [5]. In particular, we analyzed the data with an aim to identify the most representative microbial and metabolic features of T2D gut microbiota, and to gain insights into relevant microbe-microbe and microbe-host relationships.

For subsequent analyses, we selected T2D and non-diabetic control microbiome samples from a demographic cohort characterized as male, middle-aged, and normal weight. For this male, middle-aged, and normal weight cohort, we then found microbial entities differentially abundant or scarce in T2D. Here, a microbial entity represents a single microbial species or a group of multiple microbial species; a group pertains to microbial species of either a genus or a metabolic clique, which is defined here as a group of species that import, export, or degrade the same metabolite or macromolecule (e.g., glucose importers, butyrate exporters, or cellulose degraders). We then applied NJS16 as a reference map to build the community structure of microbial entities abundant or scarce in T2D, and to understand how they metabolically influence each other. This network, which we will call henceforth the microbial metabolic influence network (MIN), is based on actual microbial relative abundance information from the relevant T2D and control microbiome samples. A brief description of how we constructed these context-specific microbial MINs is as follows: one microbial entity can provide nutrients to another microbial entity via interspecies cross-feeding of metabolic byproducts or release of macromolecule degradation products. This positive impact may promote microbial growth. In contrast, a microbial entity can limit another entity's access to nutrients via competition for the same metabolites. This negative impact may inhibit microbial growth. Based on the combination of such positive and negative effects through NJS16, we formulate and quantify the net metabolic influence of one microbial entity on another microbial entity. This approach allows us to construct a community-level network of positive and negative metabolic influences between pairs of microbial entities differentially abundant or scarce in T2D. Furthermore, we include in these networks significant cross-feeding interactions between microbes and the host.

The resulting influence networks prompted us to explore whether a hidden hierarchy of metabolic influence exists within the microbial communities. To this end, we computed the T2D-relevant, community metabolic influence of each microbial entity, and obtained the cumulative number of community members toward which the microbial entity exerts a very positive or negative metabolic influence in either a direct or indirect way. Based on these numbers, we identified a set of highly interactive microbial entities, possibly acting as driving forces behind the global dynamics of their respective communities. We call these entities the network influencers. Among the total microbial entities, 17.6% were identified as network influencers in our study.

In addition, metabolites that often originate from microbial entities abundant in T2D may be indicative of how gut microbes, through their metabolism, play key roles in a specific disease context. In this regard, we sought to identify the metabolites that are most commonly produced by microbial entities abundant in either T2D or the control. From all microbial entities differentially

abundant in the non-diabetic control, the top five commonly produced metabolites (in terms of the fraction of different entities that produce a given metabolite) were acetate (77.8%), CO_2 (55.6%), lactate (55.6%), propionate (44.4%), and butyrate (44.4%) (Fig. 1, bottom left panel). Butyrate and propionate have been shown to exert multiple beneficial effects on host physiology, including intestinal glucose production (IGN, which helps prevent deregulation of glucose homeostasis and weight-gain) [6] and anti-inflammatory effects [7]. On the other hand, the top five metabolites (in terms of the fraction of different entities that produce a given metabolite) produced by microbial entities differentially abundant in T2D were acetate (25.6%), CO_2 (21.4%), ammonia (NH_3) (14.5%), hydrogen sulfide (H_2S) (14.5%), and ethanol (10.3%) (Fig. 1, bottom right panel). Acetate and CO_2 overlap with those in the aforementioned case of non-diabetic control. For the remaining metabolites unique to T2D (i.e., NH_3, H_2S, and ethanol), it warrants mentioning that NH_3 and H_2S, often the outcomes of protein fermentation processes by intestinal bacteria, are known to cause adverse health effects as carcinogenic and genotoxic agents [8]. In the context of T2D pathology, it may be worthwhile to pursue these metabolites as part of future investigations into the mechanistic relationships between gut microbial metabolic processes and T2D.

FUTURE PERSPECTIVES

To provide a global framework for understanding community metabolism within the human gut, we have constructed NJS16, a network architecture encompassing the myriad relationships among gut microbial species, host cells, and chemical compounds [3]. Each individual link in our network is from literature evidence, yet further experimental data are necessary to validate and update the global connectivity of the proposed network structure. Thus, our work calls for the need to develop high-throughput, quantitative techniques for identifying and validating specific functions (e.g., import and export of metabolites, and release of public goods from macromolecule degradation) and microbial metabolic interactions (e.g., cross-feeding mechanisms and positive/negative metabolic influences) on a global scale, specifically within *in vivo* environments. Improvements in single-cell genomics and metabolomics strategies, in culturing techniques for previously uncultured microbes, and in platforms for *in vivo* high-throughput screenings will undoubtedly accelerate this process. The advent of such technologies could confirm or update our findings, as well as push forward and establish general concepts and theories of ecological systems biology. Clearly, a comprehensive understanding of the metabolic relationships between gut microbes, and of how those relationships are intertwined with host physiology, is essential for the development of microbiota-based treatments for disease. We see our work as an important step toward this direction, by providing a computational platform for the rational design of microbial communities to benefit our health and thereby helping develop personalized clinical strategies.

Acknowledgements: I thank Jaeyun Sung, Seunghyeon Kim, Josephine Jill T. Cabatbat, Sungho Jang, Yong-Su Jin, Gyoo Yeol Jung, and Nicholas Chia for their contributions to this research. This work was mainly supported by the National Research Foundation of Korea (NRF) Grant NRF-2015R1C1A1A02037045.

References

[1] R. Sender, S. Fuchs, R. Milo, PLoS Biol. 14. e1002533 (2016).
[2] I. Sekirov, S. L. Russell, L. C. Antunes, B. B. Finlay, Physiol. Rev. 90. 859 (2010).
[3] J. Sung, S. Kim, J. J. T. Cabatbat, S. Jang, Y.-S. Jin, G. Y. Jung, N. Chia, P.-J. Kim, Nat. Commun. 6. 15393 (2017).
[4] World Health Organization. Diabetes. Geneva, Switzerland: World Health Organization (2015).
[5] J. Qin et al., Nature 490. 55 (2012).
[6] F. De Vadder et al., Cell 156. 84 (2014).
[7] N. Arpaia et al., Nature 504. 451 (2013).
[8] R. Hughes, E. A. Magee, S. Bingham, Curr. Issues Intest. Microbiol. 1. 51 (2000).

Pan-Jun Kim is a research assistant professor in the Department of Physics at the Korea Advanced Institute of Science and Technology (KAIST). He was a leader of a Junior Research Group at the Asia Pacific Center for Theoretical Physics (APCTP) and an adjunct professor in the Department of Physics at Pohang University of Science and Technology (POSTECH). He received the Young Physicist Award from the Korean Physical Society, and worked as an Institute for Genomic Biology Fellow at the University of Illinois at Urbana-Champaign after earning his PhD in statistical physics from KAIST.