– Podium Abstracts –

The Human Gut Microbiome: From Basic Science Towards Medical Applications

Peer Bork

EMBL, Heidelberg, Germany. bork@embl.de

Our knowledge of the human gut microbiome and its role for human health and well-being are constantly increasing. Yet, we don't really know how a healthy gut microbiome should look under which conditions and how it is shaped under normal circumstances during life time. In respect to the latter, I will report on selective seeding of the infant microbiome at birth and microbial transmission from the environment later in life. Regardless of the limited baselines in terms of healthy variation in relation to lifestyle, age and genetics, an increasing number of diseases is being associated with dysbiosis of the microbiome. I will illustrate the respective diagnostic potential using microbial markers for colorectal cancer as an example (Zeller et al., Mol.Sys.Biol., 2014). If dysbiosis of the microbiota causes or enhances disease, there is also a potential for therapy. Furthermore, exploitation of microbiota-drug interactions might lead to improved treatment regimes. Since the observation that the drug metformin causes a considerable shift in the gut microbiota that was formerly attributed to type 2 diabetes (Forslund et al., Nature 2015), more and more drugs are reported to perturb our gut microbiome. To study this more systematically we conducted a large in vitro screen of >1000 marketed drugs against 40 gut bacteria and found that at least 24% of all drugs directed against human cells inhibit at least one gut microbial strain. The results have multiple implications for improving drug therapies: from drug repurposing via improving the mode of action to the reduction of side effects. One successful therapy independent of drugs is faecal microbiota transplantation that performs remarkably well in patients with recurrent *C. difficile* infections, but appears to work less efficient for other conditions such as ulcerative colitis. Using single nucleotide variation (Schloissnig et al., Nature 2013, Zhu et al., Genome Biol. 2015) we show how donor strains can colonise for at least three months in the recipient, often in coexistence with indigenous strains (Li et al., Science 2016). This implies that, in principle, undesired strains could be diluted and outcompeted, opening up opportunities for novel, and personalise treatments, in analogy to our results on drug perturbations.

Robust and Reproducible Metagenomics NGS Data Analysis in the Era of the Microbiome

Patrick Chain

Los Alamos National Lab, Los Alamos, NM. pchain@lanl.gov

Metagenomics has transformed the fields of microbial ecology and microbiology, allowing us to peer at the vast and previously unknown diversity in various microbial communities. While the scientific community has been adopting shotgun metagenomics with open arms, next generation sequencing continues to pose a number of challenges when geared toward general descriptions of the communities under study. One challenge is simply how best to analyze the resulting data. Assembly has been invoked successfully to reconstruct some of the dominant members of a population, and provides a detailed overview for those communities and their potential functioning. For less dominant members of the population and for highly complex communities, the use of read-based analyses becomes of greater importance, yet is fraught with many issues including how to robustly assign reads to taxa, given an incomplete database, and short reads. Other challenges include the high computational cost associated with performing in-depth analyses that include all available genomes and/or metagenomes. Lastly, the comparisons among communities has proven to be highly effective in associating organisms and functions with different environmental parameters, however the comparison of microbiomes between studies is made more complex due to the lack of standardization of methods for data analyses. Some of these issues will be discussed in this presentation, including a series of efforts designed to lower the barrier for non-experts to use NGS for routine (meta)genomics applications by developing a user-friendly, web-based platform with a suite of bioinformatics tools.

Models for Analyzing Longitudinal Microbiota Data

Lawrence David

Duke University, Durham, NC. lawrence.david@duke.edu

Modeling how microbial communities behave over time is essential for understanding their fundamental biology, as well as for managing these communities' structure and function. Obstacles to dynamical modeling of human gut microbiota have included the relative nature of sequencing-based surveys, technical and biological uncertainty associated with measurements, and sampling limitations in generating time-series data. Here, we develop dynamical linear models that are designed for microbiota datasets and explicitly model the biological variation and technical uncertainty within them. We apply these linear models to human gut microbial communities tracked on an hourly basis for days. Our modeling reveals sub-daily dynamics to human gut microbiota, suggests guidelines for choosing optimal sampling frequencies, and provides natural links to control theory.

From Illuminating “Dark Matter” to Modelling the Microbiome

Bas E. Dutilh

Assistant Professor, Theoretical Biology and Bioinformatics, Utrecht University, Utrecht, The Netherlands.
And, Centre for Molecular and Biomolecular Informatics, Radboud University Medical Centre, Nijmegen, The Netherlands. bedutilh@gmail.com

Metagenomics has unveiled a vast microbial biodiversity in a range of environments. Beyond inferring the distribution of species and functions in microbiomes, bioinformatic advances now allow complete microbial and viral genomes to be reconstructed with an unprecedented resolution. This genomic resolution invites systems-level modelling, towards a mechanistic understanding of the role of the different species within their community. I will show some examples of how genome-resolved metagenomics can be used to predict interactions, and provide initial models of the microbial ecosystem in various human body sites.

Studying Genomic Heterogeneity Using Single-Amino Acid Variants

A. Murat Eren

Assistant Professor, University of Chicago, Chicago, USA. meren@uchicago.edu

The diversity and geographical distribution of populations within major marine microbial lineages are largely governed by temperature and its co-variables. However, neither the mechanisms by which genomic heterogeneity emerges within a single population nor how it drives the partitioning of ecological niches are well understood. We took advantage of billions of metagenomic reads from the TARA Oceans project to study one of the most abundant and widespread microbial SAR11 populations in the surface ocean, and characterized its substantial amount of genomic heterogeneity using "single amino-acid variants" (SAAVs). We identified systematic purifying selection and adaptive mechanisms governing non-synonymous variation within this population that revealed two broad ecological niches that reflect large-scale oceanic current temperatures, as well as six proteotypes that delineate more subtle niche demarcations. We also identified significantly more protein variants in cold currents and an increased number of protein sweeps in warm currents, exposing a global pattern of alternating genomic diversity for this SAR11 population as it drifts along with surface ocean currents. Overall, our results show that purifying selection constrains the scope of neutral evolution to amino acid sequence variants permitted by protein stability requirements, and the geographic partitioning of SAAVs suggests natural selection, rather than neutral evolution, is the main driver of the evolution of SAR11 in surface oceans.

Assessing Potential Interactions Among Marine Bacteria, Archaea, Protists, and Viruses via 'Omics and High Resolution Time Series

Jed Fuhrman

McCulloch-Crosby Chair of Marine Biology
University of Southern California, Los Angeles, CA, USA. fuhrman@usc.edu

Due to the difficulty of directly observing the largely-invisible activities of marine microorganisms in their natural habitat, we have developed molecular genetic tools to examine the abundances and activities of a broad array of microbes, ranging from viruses up to protists, in high resolution time series to begin evaluating such interactions. To our surprise, communities can change remarkably rapidly, in ways different from what the textbooks say. We use network analyses to evaluate microbial associations, learning new such associations and additional details about known ones. One of the hardest aspects is determining which 'omics' sequences are from "unknown" viruses and to then link such viruses to possible hosts; we have worked with computational biologists to develop some sophisticated adjusted k-mer profile-based methods to address these tasks with some success. We also find that metatranscriptomes provide particularly useful "smoking guns" to show viral activity.

From Microbiomes to Resistomes: Gut Microbial Community Dynamics in Chronically Antibiotic-Treated Patients

Uri Gophna

Professor of Biotechnology, Tel Aviv University, Israel. microevol@gmail.com

Pouchitis is a common long-term complication in patients with ulcerative colitis undergoing proctocolectomy with ileal pouch-anal anastomosis. Patients may become chronically-treated with antibiotics in order to maintain disease remission. This raises the question of how gut microbial communities respond to such long-term selective pressures, in terms of stability and evolution of resistance. We performed shotgun metagenomics on longitudinal samples from antibiotic-treated patients and controls treated by different drugs, such as anti-TNF monoclonal antibodies and steroids. We then analyzed these metagenomes combining resistance database data with alignment-guided prediction of quinolone resistance gene alleles. Opportunistic and commensal members of the gut microbiome of patients with antibiotic-treated pouchitis develop known antibiotic resistance mutations. Constant long-term antibiotic treatment further intensified bacterial community instability in pouchitis patients. Presence of drug-resistant *Escherichia coli* and *Enterococcus faecalis* strains was confirmed by detecting single-point mutations in two drug target genes *gyrA* and *parC* in the antibiotic-treated patients, but these strains had very few virulence genes, indicating a low pathogenic potential. Nonetheless, several bacterial species considered beneficial also showed similar resistance mutations.

The Microbial Genome Atlas (MiGA) project

Kostas Konstantinidis

Carlton S. Wilder Associate Professor, Georgia Institute of Technology, Atlanta, GA, USA.
kostas@ce.gatech.edu

The small subunit ribosomal RNA gene (16S rRNA) has been successfully used to catalogue and study the diversity of microbial species and their communities to date as exemplified by the Ribosomal Database Project (RDP; https://rdp.cme.msu.edu). Nonetheless, several aspects of rRNA gene-based studies remain problematic. Most importantly, how to better resolve microbial communities at levels at which the 16S rRNA gene provides inadequate resolution, namely the species and finer levels, and how to best catalogue whole-genome diversity and fluidity. To bridge this gap, we have developed the "genome-equivalent" of RDP called the MiGA project (available at: www.microbial-genomes.org). MiGA allows the classification and gene-content diversity study of query genome(s) or assembled contig(s) against a reference database of microbial genomes using the ANI/AAI concept (currently using the ~13,000 isolate genomes available in NCBI's Genome database). Examples of using MiGA to perform high-resolution microbial source tracking in riverine ecosystems as well as micro-diversity and epidemiological studies of bacterial pathogens will be presented. Our more recently developed computational solutions that will allow MiGA to scale-up with the increasing volume of genomic and metagenomic sequence data that are becoming available, and encompass the "uncultivated majority" of microbes, will be also highlighted.

The New Age of Virus Metagenomics

Eugene V. Koonin

Senior Investigator, National Center for Biotechnology Information, NLM, National Institutes of Health, Bethesda, Maryland, USA. koonin@ncbi.nlm.nih.gov

With the recent rapid advances of metagenomics, the great majority of new viruses that are currently discovered are only known as genomes assembled from metagenomics sequences. The role of metagenomics in virus discovery has become so prominent that the International Committee for Taxonomy of Viruses (ICTV) has formally allowed recognition of virus taxa on the basis of the genome sequence alone, which is a major departure from established practices in taxonomy. It seems that the very layout of the genomics enterprise has changed. Nowadays, when a new virus is discovered, it is nearly certain that an entire family of its relatives can be identified as well, allowing one to quickly identify the key functional elements and evolutionary relationships solely by computational methods. However, the success of this type of analysis critically depends on judicious use of the most powerful available methods for sequence and structure analysis. I will illustrate the new age of viral (meta)genomics by three examples: 1) discovery of new family of bacteriophages that apparently infect bacteria of the phylum Bacteroidetes and include crAssphage, the most abundant member of the human-associated virome; 2) discovery of a new family of archaeal viruses that appear to be major players in the world ocean ecology; 3) discovery of a new group of giant viruses of unicellular eukaryotes that sheds new light on virus evolution.

Microbiome Research: from Products to Data

Nikos Kyrpides

DOE-Joint Genome Institute, Walnut Creek, CA. nckyrpides@lbl.gov

Microbiome research is rapidly transitioning into Data Science. The unprecedented volume of microbiome data being generated pose significant challenges with respect to standards and management strategies, but also bear great new opportunities that can fuel discovery. Computational analysis of microbiome samples involving previously uncultured organisms, is currently advancing our understanding of the structure and function of entire microbial communities and expanding our knowledge of genetic and functional diversity of individual micro-organisms. I will describe some of our computational approaches and will emphasize the value of data processing integration in enabling the exploration of large metagenomic datasets and the discovery of novelty. I will discuss current approaches and success stories for the discovery of novel phylogenetic lineages as well as the exploration of the viral dark matter.

Controlling Human Microbiota

Yang-Yu Liu
Assistant Professor, Harvard Medical School, Harvard University, Cambridge, USA.
spyli@channing.harvard.edu

We coexist with a vast number of microbes—our microbiota—that live in and on our bodies, and play an important role in human physiology and diseases. Propelled by metagenomics and next-generation DNA sequencing technologies, many scientific advances have been made through the work of large-scale, consortium-driven metagenomic projects. Despite these advances, there are still many fundamental questions regarding the dynamics and control of microbiota to be addressed. Indeed, it is well established that human-associated microbes form a very complex and dynamic ecosystem, which can be altered by drastic diet change, medical interventions, and many other factors. The alterability of our microbiome offers opportunities for practical microbiome-based therapies, e.g., fecal microbiota transplantation and probiotic administration, to restore or maintain our healthy microbiota. Yet, the complex structure and dynamics of the underlying ecosystem render the quantitative study of microbiome-based therapies extremely difficult. In this talk, I will discuss our recent theoretical progress on controlling human microbiota.

There Is a World Going on Underground: Gene Flow and Recombination Among Subsurface *Thermotogae* Populations

Camilla L. Nesbø

University of Alberta and BioZone University of Toronto, Edmonton, Canada.
nesbo@ualberta.ca

I will discuss two studies where we have used comparative genomic- and phylogeographic approaches to elucidate the histories of anaerobic subsurface bacterial populations. Oil reservoirs represent nutrient-rich islands within the subsurface and most contain diverse microbial communities. How indigenous microbes come to inhabit these environments still remains uncertain; they may colonize the oil reservoirs from surrounding subsurface populations or they may have been buried with the sediments that later make up the oil reservoirs. To address this question, we have used a comparative genomic approach to investigate the phylogeographic patterns of hyperthermophilic and mesophilic *Thermotogae* bacteria common in deep and shallow oil reservoirs, respectively. For hyperthermophilic *Thermotoga maritima*-like bacteria, genome analyses revealed extensive gene flow within, but also between, subsurface and marine hydrothermal vent populations, supporting the 'colonization hypothesis' where *Thermotoga* bacteria from subsurface and marine populations have been continuously migrating into the oil reservoirs, influencing their genetic composition. Analyses of genomes of mesophilic *Mesotoga* bacteria, on the other hand, revealed differentiated lineages with low levels of genetic contact between groups. The phylogeographic patterns observed suggests that geographic separation, and possibly 'burial and isolation', have been more important for these mesophilic populations. This somewhat counterintuitive observation, where mesophiles are more affected by geography than thermophiles, is likely due to the anaerobic lifestyle of the *Thermotogae*.

Near-Real Time Personalized Medicine and Cystic Fibrosis

Forest Rohwer

San Diego State University, San Diego, CA. frohwer@gmail.com

Near real-time microbiology approaches will enable doctors to make better decisions about patient treatments. The San Diego research community has established a collaborative effort to generate and interpret metagenomics, metatranscriptomic and metabolomics (i.e., -omics) data from cystic fibrosis (CF) sputum samples in approximately 24 hours. This work is part of a greater background of a long-term sampling effort, where each patient serves as their own benchmarks for different disease states. This approach allows us to more rapidly determine what has changed at any particular time in the patient's history. Using these "-omics" data we are identifying the underlying viral and microbial mechanisms that drive the cyclical nature, stable, exacerbation and recovery, of CF. This background data is extremely useful for diagnosing what is unique about fatal exacerbations and points to possible treatment options. A case study in which we applied this approach to a patient experiencing a fatal exacerbation event will be presented. The "-omics" techniques complimented each other and allowed us to determine the most likely cause of the event. I will also discuss what learned that worked versus what did not work.

Personalizing Treatments Using Microbiome and Clinical Data

Eran Segal

Professor, Department of Computer Science and Applied Mathematics,
Weizmann Institute of Science, Israel. Eran.Segal@weizmann.ac.il

Accumulating evidence supports a causal role for the human gut microbiome in obesity, diabetes, metabolic disorders, cardiovascular disease, and numerous other conditions, including cancer. Here, I will present our research on the role of the human microbiome in health and disease, aimed at developing personalized medicine approaches that combine human genetics, microbiome, and nutrition.

In one project, we set out to understand personal variation in the glycemic response to food, tackling the subject of personalization of human nutrition, a poorly studied topic that is critical for human health and to billions of people predisposed to, or suffering from, obesity, T2D and related co-morbidities. We assembled a 1,000 person cohort and measured blood glucose response to >50,000 meals, lifestyle, medical and food frequency questionnaires, blood tests, genetics, and gut microbiome. We showed that blood glucose responses to meals greatly vary between people even when consuming identical foods; devised the first algorithm for accurately predicting personalized glucose responses to food based on clinical and microbiome data; and showed that personalized diets based on our algorithm successfully balanced blood glucose levels in prediabetic individuals. These results suggest that personalized diets may successfully modify elevated postprandial blood glucose and its metabolic consequences.

I will also present our studies of the mechanisms driving recurrent post-dieting obesity in which we identified an intestinal microbiome signature that persists after successful dieting of obese mice. This microbiome signature contributes to faster weight regain and metabolic aberrations upon re-exposure to obesity-promoting conditions and transmits the accelerated weight regain phenotype upon inter-animal transfer. These results thus highlight a possible microbiome contribution to accelerated post-dieting weight regain, and suggest that microbiome-targeting approaches may help to diagnose and treat this common disorder.

Finally, we studied the relative contribution of host genetics and environmental factors in shaping human gut microbiome composition. To this end, we examined genotype and microbiome data in over 1,000 healthy individuals from several distinct ancestral origins who share a relatively common environment, and demonstrated that the gut microbiome is not significantly associated with genetic ancestry. In contrast, we find significant similarities in the microbiome composition of genetically unrelated individuals who share a household, and show that over 20% of the gut microbiome variance can be explained via environmental factors related to diet, drugs and anthropometric measurements. We define the term biome-explainability as the variance of a host phenotype explained by the microbiome after accounting for the contribution of human genetics. Consistent with our finding that microbiome and host genetics are largely independent, we find significant biome-explainability levels of 24%-36% for several human traits and disease risk factors. We also successfully replicated our results in an independent Dutch cohort. Overall, our results suggest that human microbiome composition is dominated by environmental factors rather than by host genetics.

Human Skin Microbiome: Topographic Functional Mapping of Healthy Volunteers and Patient Populations

Julia A. Segre[1], Allyson L. Byrd[1], Osnat Tirosh[1], Heidi H. Kong[2]

[1]Chief & Senior Investigator, National Human Genome Research Institute; [2]National Institute of Arthritis, Musculoskeletal and Skin Diseases, NIH, Bethesda, Maryland, USA.
jsegre@nhgri.nih.gov

The varied topography of human skin offers a unique opportunity to study how the body's microenvironments influence the functional and taxonomic composition of microbial communities. The general accessibility of skin also enables longitudinal clinical studies throughout the course of disease manifestations. Metagenomic analysis of diverse body sites in healthy humans defined the skin microbiome as shaped by the local biogeography, yet marked by strong individuality. We further identified strain-level variation of dominant species as heterogeneous and multiphyletic. Re-sampling months and years later revealed that despite the skin's exposure to the external environment, its bacterial, fungal, and viral communities were largely stable over time. Strain and single nucleotide variant level analysis showed that individuals maintain, rather than reacquire prevalent microbes from the environment.

Longitudinal stability of skin microbial communities generates hypotheses about colonization resistance and empowers clinical studies exploring alterations observed in disease states, such as the inflammatory skin disorder atopic dermatitis (AD; commonly known as eczema). Integrating shotgun metagenomic sequencing, culturing, and animal models, we explore a model whereby staphylococcal strains contribute to AD disease exacerbation. This delineation of highly individualized skin microbiomes with patient-specific strains underscores the individuality of the disease course and therapeutic response in AD patients and may represent an opportunity for precision medicine.

Although landmark studies have shown that microbiota activate and educate host immunity, how the immune system shapes microbial communities and contributes to disease is less-well characterized. We explore the skin microbiome of patients with primary immune deficiencies to address this question and to expand our understanding of microbes that colonize human skin.

Understanding Viruses in Nature May Save the Earth and Cure Your Next Ailment

Matthew Sullivan

Associate Professor, Ohio State University, Columbus, OH. sullivan.948@osu.edu

Microbes are recently recognized as driving the energy and nutrient transformations that fuel Earth's ecosystems in soils, oceans and humans. Where studied, viruses appear to modulate these microbial impacts in ways ranging from mortality and nutrient recycling to extensive metabolic reprogramming during infection. As environmental virology strives to get a handle on the global virosphere (the diversity of viruses in nature), we face challenges to organize this 'sequence space' (create a sequence-based viral taxonomy), link these viruses to their natural hosts (who infects whom), and establish how virus populations are structured (ecological drivers) and impact natural ecosystems (their impacts). Here I will share current thinking on how to study viruses in ocean communities and how these efforts are now revealing new biology in the clinic.

A Null Model for Microbial Diversification

Olga Zhaxybayeva

Assistant Professor, Dartmouth College, Hanover, NH. Olga.Zhaxybayeva@dartmouth.edu

One of the first tasks in an analysis of a microbial metagenome is to assign obtained DNA fragments to their organismal source ("species"). However, microbiologists do not yet agree on a coherent species concept for Bacteria and Archaea. Without such a concept, metagenomic studies employ *ad hoc* conventions for identification and quantification of prokaryotic diversity. For example, a commonly used approach is to group sequences into "operational taxonomic units" delimited by regions of 16S ribosomal RNA genes showing at least 97% sequence identity. Such conventions are based on observations that gene sequences from closely related prokaryotic groups are typically organized into distinct clusters. However, we do not know if these clusters represent fundamental units of bacterial diversity ("species"), nor the precise nature of evolutionary and ecological forces that are responsible for cluster formation. Multiple processes, such as periodic selection and extensive recombination, have been proposed to be drivers of cluster formation ("speciation"). Yet, the clustering patterns are rarely compared to those obtainable with a simple null model of diversification under stochastic lineage birth and death and random genetic drift. In my talk, I will show that this simple model could also explain the observed clustering. I will argue that testing for the signatures of such "neutral" patterns should be considered a null hypothesis in any microbial classification analysis. And it is only when the real data are statistically different from the expectations under the null model that some speciation process should be invoked.