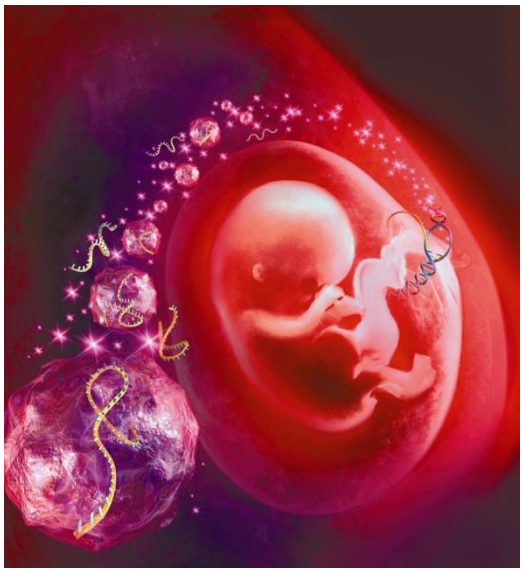
King Faisal
INTERNATIONAL PRIZE



THE KING FAISAL MEMORIAL ARTICLES IN MEDICINE AND SCIENCE XIIV

THE 2015 KING FAISAL INTERNATIONAL PRIZE

THE KING FAISAL MEMORIAL ARTICLES IN MEDICINE AND SCIENCE XIIV

THE 2015 KING FAISAL INTERNATIONAL PRIZE

The King Faisal International Prize
Post Box 22476, Riyadh 11495
Saudi Arabia

Custodian of the Two Holy Mosques
King Salman Ibn Abd al-Aziz Al-Saud
Patron of the King Faisal Foundation

HRH Prince Muqrin ibn Abd Al-Aziz Al-Saud
Crown Prince and First Deputy Prime Minister

HRH Prince Muhammad bin Nayef bin Abdulaziz Al Saud
Deputy Crown Prince and
Second Deputy Prime Minister
Minister of Interior

Contents

INTRODUCTION

The King Faisal Foundation continues the traditions of Arabic and Islamic philanthropy, as they were revitalized in modern times by King Faisal. The life and work of the late King Faisal bin Abd Al-Aziz, son of Saudi Arabia's founder and the Kingdom's third monarch, were commemorated by his sons and daughters through the establishment of the Foundation in 1976, the year following his death. Of the many philanthropic activities of the Foundation, the inception of King Faisal International Prizes for Medicine in 1981 and for Science in 1982 will be of particular interest to the reader of this book. These prizes were modeled on prizes for Service to Islam, Islamic Studies and Arabic Literature which were established in 1977. At present, the Prize in each of the five categories consists of a certificate summarizing the laureate's work that is hand-written in Diwani calligraphy; a commemorative 24-carat, 200 gram gold medal, uniquely cast for each Prize and bearing the likeness of the late King Faisal; and a cash endowment of SR750,000 (US$200,000). Co-winners in any category share the monetary award. The Prizes are awarded during a ceremony in Riyadh, Saudi Arabia, under the auspices of the Custodian of the Two Holy Mosques, the King of Saudi Arabia or his representative.

Nominations for the Prizes are accepted from academic institutions, research centers, professional organizations and other learned circles worldwide, as well as from previous laureatues. After preselection by expert reviewers, the short-listed works are submitted for further, detailed evaluation by carefully selected international referees. Autonomous, international specialist selection committees are then convened at the headquarters of the King Faisal Foundation in Riyadh each year in January to make the final decisions. The selections are based solely on merit, earning the King Faisal International Prize the distinction of being among the most prestigious of international awards to physicians and scientists who have made exceptionally outstanding advances which benefit all of humanity.

WINNERS OF THE 2015
KING FAISAL INTERNATIONAL PRIZE
FOR MEDICINE

The King Faisal International Prize for Medicine for this year 2015G (1436H), Topic: Intestinal Microflora and Human Health, has been awarded to: Professor Jeffrey Ivan Gordon (USA) Dr. Robert Glaser Distinguished University Professor, and Director of the Centre of Genome Sciences and Systems Biology at Washington University, St Louis, USA. In recognition of his seminal work on defining the microbiomes genomic and metabolic role in human health. Professor Gordon's pioneering work and interdisciplinary studies of the human microbiome has provided fascinating insights into the metabolic processes and the genetic basis of mutually beneficial relationships between the host and microorganisms in the human gut. His innovative research has provided major breakthroughs into the influence of intestinal microbiota on postnatal development, physiology and illness susceptibility in humans and has thus enhanced our understanding of the pathogenesis of complex diseases such as obesity. His research has opened opportunities for novel gut microbiome directed treatments to improve human health.

A journey together: exploring the human gut microbiota

Jeffrey I. Gordon

Center for Genome Sciences and Systems Biology, and Center for Gut Microbiome and Nutrition Research, Washington University School of Medicine, St. Louis, MO 63018

If you want to go fast, go alone. If you want to go far, go together.
African proverb

In looking back to recount the lab's journey, I feel today what I have always felt; that I have been incredibly fortunate to have been surrounded by an inspiring group of remarkable students with enormous curiosity, a deep commitment to learning, the courage to explore the unknown, and a willingness to cross traditional disciplinary boundaries to obtain answers to questions they found to be captivating. They have had a wonderful sense of community - a shared belief that discoveries and innovation are born from a caring, supportive respectful and trusting environment where we can communicate our ideas freely with one another and not be afraid to say 'I don't understand' so that we learn together. They have manifested that wonderful mix of idealism, imagination and other directedness that allows people to look beyond themselves to embrace and together pursue with hope, modesty and humility, using both basic and applied science, solutions to challenging global health problems that are necessary to address now rather than later and for sustained periods of time.

Our journey began with a series of questions related continuously renewing epithelial lining of the mouse and human gut. How do each of its four continuously renewing epithelial cell lineages establish and maintain their different spatial patterns of gene expression along the length of the intestine: e.g., why do members of the enterocytic lineage express different sets of gene products in the duodenum compared to the ileum or colon? What factors are responsible for instructing these epithelial cells to execute developmental stage-specific differences in their gene expression before and after birth? What were the environmental cues that shape regional and temporal patterns of cellular differentiation?

The approach we initially adopted was to use transgenic mice to map cis-acting transcriptional regulatory elements that controlled the cell lineage-specific, developmental stage-specific and spatial patterns of expression of a group of prominently expressed genes involved in the binding and trafficking of lipids (e.g., Sweetser et al., 1988). Over time, we assembled a 'toolbox' of these regulatory elements that allowed us to deliver, in genetically engineered mice, foreign gene products to different epithelial cell lineages located in different regions along the length of the gut, at various times during pre- and postnatal development, and at specified points during their differentiation. This differentiation is executed during an orderly migration from the base of crypts of Lieberkuhn, where the lineages cells originated from multipotential stem cells, to the tips of small intestinal villi, or the colonic homolog of small intestinal villi, the surface epithelial cuff.

This toolbox allowed Michelle Hermiston, a MD/PhD student, to generate chimeric mice composed of the descendants of genetically manipulated embryonic stem (ES) cells and the descendants of non-manipulated blastocyst-derived cells (Hermiston et al, 1993, Hermiston and Gordon, 1995; Hermiston et al., 1996). Each finger-like intestinal villus is supplied by several crypts, which continuously feed epithelial cells to that villus. The epithelial cell population in each crypt is monoclonal, arising from a single active stem cell. In a chimeric mouse, a given villus can be fed by monoclonal crypts of ES cell or blastocyst origin; the products of each crypt migrate up a given villus in coherent, sharply demarcated columns composed of differentiating epithelial cells. Once these cells approach they villus tip, they are removed (undergo apoptosis; exfoliate into the intestinal lumen). The process is perpetual, with three of the four epithelial cell lineages (enterocyte, globlet and enteroendocrine) renewing themselves approximately once a week in the mouse and every two weeks in humans (members of the Paneth cell lineage migrates downward to the crypt base and are longer-lived) (Gordon, 1989). A villus fed by ES and blastocyst derived crypts

represents a well controlled experiment: the juxtaposed columns of genetically manipulated ES-derived epithelium and non-manipulated blastocyst-derived epithelium allow the effects of the genetic manipulation to be ascertained at a given location along the crypt-villus axis in a given region of the small intestine. Genetically slowing the rate of migration of ES-derived epithelial cells along the crypt-villus axis through forced expression of E-cadherin illustrated to us how epithelial differentiation was critically dependent on positional cues. Slowing migration did not result in the expression of gene products earlier in the migration pathway; rather, cells had to reach a given location in their journey along the villus to induce expression of differentiation-associated gene products (Hermiston et al., 1996).

These results spawned a debate in the lab; where to look for the source of instructions that regulate various facets of cellular differentiation? As developmental biologists, we were naturally tempted to turn to the underlying mesenchyme given the general importance of epithelial-mesenchymal cross talk in many systems. However, thinking about intestinal development, we could not ignore the fact that beginning at birth, it is colonized by a vast microbial community (microbiota). We reasoned that the ability of members of the microbiota to establish themselves along the length of the gut could reflect a set of reciprocal signaling events where the epithelium would provide nutritional resources that could be harvested by early colonizers. These colonizers in turn could influence gene host expression in ways that would allow them, as well as other organisms, to establish residency in a given region of the gut. We imagined these reciprocal interactions to progress influencing host and microbial community characteristics in ways that would allow continued microbial succession in a dynamically evolving gut habitat. The underlying assumption (or should I say, hope) was that there was a discernible order to this process, and despite the enormous apparent complexity of the system (the number of potential combinatorial microbe-microbe and microbe-host interactions is mind-boggling), it would be possible to dissect these interactions using simplified models of the human gut ecosystem.

The debate continued about how to create a simplified system composed of human microbial parts. We decided to start with a human gut symbiont that could be genetically manipulated, an anaerobe that wasn't difficult to culture, and one that exemplified a key functional feature of the microbiota – the capacity to break down complex polysaccharides that were not digestible by humans because our genome lacked the glycoside hydrolases and polysaccharide lyases needed to do so. We traveled 180 miles north from St. Louis to Abigail Saylers' lab at the University of Illinois in Urbana-Champaign and asked her about an

organism with a barely pronounceable name that she had been studying, *Bacteroides thetaiotaomicron*. She had developed genetic tools and had characterized its ability to break down a range of dietary and host polysaccharides *in vitro*. With characteristic generosity, she opened her freezer (literally and metaphorically) and gave us all that we needed to begin.

We were captivated by gnotobiotics: the ability to rear mice under germ-free conditions and then to colonize them at a selected time during their lives with a given organism or group of microbes. Gnotobiotic mouse models seemed like the perfect way to orchestrate (stage), in a deliberate highly controlled way, interactions between a human gut symbiont or symbionts and a mammalian host. Lynn Bry, a courageous MD/PhD student, with the help of a Per Falk, a Swedish post-doc in the lab who had been studying spatial differentiation of the gut epithelium using a variety of lectins to characterize regional variations in glycan production, made arrangements to travel to the gnotobiotic mouse facility at the Karolinska Institutet run by Tore Midtvedt, one of the fathers of gnotobiology. She would do experiments there since we didn't have our own gnotobiotic facility and then bring the tissues back to St. Louis to perform her analyses.

Her results were striking and profoundly influenced the course the lab would take. Germ-free mice initiated a program of expression of fucosylated glycans in the goblet cell and enterocytic lineages located in the distal small intestine: induction occurred during the suckling period (**Figure 1A-H**). Unlikely conventionally raised mice (animals that had been exposed to microbes in their environment beginning at birth), this program of fucosylated glycan production did not generalize in germ-free animals to involve more epithelial cells over a greater region of the distal small intestine. However, colonization of weaned germ-free mice with the fecal microbiota of a conventionally raised mouse (a process known as 'conventionalization') restarted this arrested program of expanding fucosylated glycan production. Moreover, colonization of weaned germ-free mice with a wild-type strain of *B. thetaiotaomicron* alone was sufficient to direct this region-specific epithelial fucosylation. Strikingly, experiments using an isogenic mutant strain of *B. thetaiotaoimcron* that was unable to utilize fucose disclosed that this induction was dependent upon the ability of the organism to use L-fucose as a carbon source (Bry et al., 1996) (**Figure 1I,J**).

Lynn's results were provocative and inspiring at many different levels. The idea that a gut microbe could direct its host to manufacture a nutrient source that the microbe in turn could use was captivating. It also pointed to a future in which we might be able to deliberately build up complexity by colonizing gnotobiotic mice with artificial human gut

communities composed of two, three, four or more cultured members of the microbiota. These artificial communities could provide answers to basic questions concerning how microbes cooperate with one another, how they compete with one another, what are the determinants of their success in different nutrient contexts, and how each member singly or collectively shapes host biology.

We again turned to Abigail Salyers (and Tore Midvedt) for advice about establishing our own gnotobiotic facility. We also asked William Peck, Dean of our medical school, for permission and funds to transform 1000 square feet of an unfinished floor in a newly built animal facility to one open room where we could rear germ-free animals in flexible plastic film gnotobiotic isolators. He listened to our scientific plans, looked at our data, understood our dreams, and said he thought we would need more space than we had asked for. In a sentence that is rarely heard in most academic medical centers, he asked if we would mind if he provided resources to build a facility twice as large as we had originally envisioned. He was right about our future needs, and we were awestruck by our good fortunate to be in such a supportive institution.

We have also been incredibly fortunate to have two people in the lab, Dave O'Donnell and Maria Karlsson, who have overseen this facility since its birth. They have taught students how to perform their own gnotobiotic experiments. Their constant innovations allowed us to do more to characterize these representations of microbe-host interactions in health and disease. Their legacy has been several generations of students who have gone on to establish their own labs and their own gnotobiotic facilities. Dave and Maria have always been available to them after they left the lab so that we can continue to learn together as we advance our understanding and knowledge of the opportunities that gnotobiotics present.

Inspired by Lynn's results, Lora Hooper, a gifted post-doc in the lab with a backbground in glycobiology, performed a genetic analysis of *B. thetaiotaomicron* and how its fucose utilization system regulated signaling to the host (Hooper et al., 1999). She also undertook a more comprehensive functional genomics study based on use of laser capture microdissection of the gut epithelium of gnotobiotic mice colonized with *B. thetaiotaomicron* and newly introduced mouse GeneChips (Hooper et al., 2001a). This marriage of gnotobiotics and functional genomics revealed a number of unanticipated host responses and greatly expanded our appreciation of how a gut symbiont might influence many aspects of host biology. It also emphasized that if we were to understand the mutual benefit derived from establishing a microbiota, we needed to 'listen' to the host-microbial dialogue armed with better knowledge of the microbial genome.

B. thetaiotaomicron was not a pathogen. It was 1999 and large genome sequencing centers were engaged in a massive effort to define the genes in our *H. sapiens* genome. They were reluctant to divert efforts to define the genes present in members of our indigenous microbial communities (the microbiome). In 1999, students in the lab led by Jian Xu decided to do tackle the job of sequencing the *B. thetaiotaomicron* genome themselves. We obtained a grant from the Swedish pharmaceutical company Astra, which was deeply involved in studying gut health. Their resources allowed us to purchase an ABI 3700 DNA sequencer and then to sequence, assemble (finish) and annotate the bacterial genome (Xu et al., 2003). It was a critical turning point for us; we had installed the experimental and computational tools needed for comparative genomics in the lab. What we saw was amazing: *B. thetaiotaomicron* alone had more genes encoding glycoside hydrolases and polysaccharide lyases that in our human genome. These bacterial genes were organized into polysaccharide utilization loci (PULs), each of which had its own transcriptional regulator. It appeared that this organism was very well equipped to detect, import, and process a variety of host as well as dietary polysaccharides (subsequently established experimentally in a series of followup studies by a number of individuals in the lab, including Justin Sonnenburg, Eric Martens and Nate McNulty, through a combination of genetic, functional genomic and biochemical analyses conducted *in vitro* and *in vivo*; Sonnenburg et al., 2005; Martens et al., 2008, 2009, 2011; McNulty et al., 2013).

We expanded our efforts to sequence more members of the human gut microbiota. Encouraged by the wealth of information about the functional capabilities of these symbionts emanating from these early sequencing efforts, we extended our earlier suggestion for a human microbiome 'genome anatomy project' (Hooper et al., 2001b) to a formal request, in the form of a 'white paper', to the National Human Genome Research Institute to expand their efforts to define our human genetic landscape by launching a project to sequence the genomes of 100 cultured, phylogenetically diverse members of the human gut microbiota. These sequenced genomes could serve as a reference for interpreting the results of subsequent 'shotgun sequencing' of human gut microbial microbial community DNA (http://www.genome.gov/pages/research/sequencing/seqproposals/hgmiseq.pdf).

Fredrik Backhed, a newly arrived insightful postdoc, wanted to expand our view of microbiota function by examining the relationship between the microbiota, diet and obesity. Reasoning that mammalian gut microbial communities provide a survival advantage to their hosts by promoting physiologic and metabolic adaptations to sustained periods of

nutrient deprivation as well as to intermittent periods of nutrient excess, he hypothesized that it should be considered as a component of our energy balance equation. Corollaries to this hypothesis are that diet and the structural and functional configurations of a microbiota are intimately and dynamically interrelated, and that the nutritional values of foods are not absolute terms but rather are influenced in a biologically significant way by the structure and operations of a consumer's gut community.

Fredrik began by studying wild-type mice and those with a genetically engineered deficiency of a circulating inhibitor of lipoprotein lipase (Fiaf/Angpt4) that is normally secreted by the gut epithelium. Comparing wild-type adult germ-free, conventionally raised, and conventionalized mice, he showed that the gut microbiota regulates adiposity (germ-free animals are significantly leaner than their colonized counterparts). Moreover, he found that gut epithelial expression of Fiaf/Angpt4 is suppressed by the microbiota, and that this protein plays a role in mediating the effects of the microbiota on adiposity (Backhed et al., 2004). Fredrik went on to show that germ-free mice are resistant to diet-induced obesity (Backhed et al., 2007).

Ruth Ley, a post-doc with a background in ecology and environmental microbiology, used her considerable talents to direct a large bacterial 16S rRNA-based sequencing project which revealed that genetically obese, leptin-deficient *ob/ob* mice had pronounced differences in the structure of their gut (fecal) microbiota compared to their lean wild-type (+/+) or heterozygous *ob/+* littermates, including a marked reduction in the proportional representation of members of the phylum Bacteroidetes and an increase in the proportional representation of the Firmicutes (Ley et al., 2005; her comparative analysis of fecal microbiota samples obtained from phylogenetically diverse mammalian species revealed that the Firmicutes and Bacteroidetes are two dominant phyla and that diet has played a dominant role in shaping community configuration during the course of mammalian evolution; Ley et al, 2008).

Ruth's analysis of obese *ob/ob* and their lean *ob/-* and +/+ littermates gave birth to a long standing collaboration with Rob Knight, who together with his student, Cathy Lozupone, had just developed a software tool, which they called UniFrac, that compared similarities and differences between microbial communities based on the degree to which they shared branch length on a bacterial tree of life. Our interactions with Rob's lab over the years emphasized how new experimental and computational approaches co-evolve in a dynamic and interdependent manner.

Ruth translated results from her mouse model to a clinical study conducted with our colleague Sam Klein. Two groups of obese humans

were followed over the course of one year as members of each group consumed one of two types of low calorie diets: their fecal microbiota was sampled on a number of occasions as they lost weight. She found that weight loss is associated with a change in gut microbial ecology, including an increased representation of Bacteroidetes (Ley et al., 2006). This finding provided an initial link between the gut microbiota and human obesity.

Considering Koch's postulates from a different perspective, focusing not on a single isolated putative pathogen, but rather a whole community and its properties, Peter Turnbaugh, an extraordinary PhD student, transplanted gut microbiota from conventionally raised *ob/ob* mice into wild-type germ-free mouse recipients. He found that compared to microbiota from adult, conventionally raised, lean wild-type or *ob/+* mouse donors, *ob/ob*-associated microbial communities transmitted an increased adiposity phenotype. These differences were not attributable to differences in food consumption among transplant recipients but were associated with differences in microbial community metabolism (Turnbaugh et al., 2006). Peter went on to show that fecal microbiota from adult conventionally raised wild-type mice with diet-induced obesity were also capable of transmitting an increased adiposity phenotype to wild-type germ-free animals (Turnbaugh et al., 2008).

We were fascinated by the idea of turning to twins for followup human studies of the relationship between the gut microbiota and nutritional status. We reasoned that comparing monozygotic and dizygotic twin pairs concordant or stably discordant for obesity would allow us to more readily develop and test hypotheses. Potentially confounding variables related to environmental exposures and host genetics could be constrained. We could assess the generality of our findings by characterizing multiple families with twin pairs. Moreover, we could test whether a causal relationship existed between the gut microbial community configuration and health status by transplanting microbiota from each co-twin in a pair discordant for nutritional status (and from different pairs) to groups of recipient germ-free mice.

A partnership with our wonderful Washington University colleague, Andrew Heath, who had been following a large twin cohort for many years, allowed Peter to characterize the gut (fecal) microbiota of twin pairs concordant for obesity or leanness, and their mothers (Turnbaugh et al., 2009a). His metagenomic analyses revealed that (i) *intra*personal variation in gut community composition was less than *inter*personal variation (followed over time, we are our own best controls), (ii) *inter*personal variation in the bacterial species composition of the microbiota was considerably less among family members compared to unrelated individuals living in separate households (our

familial genetic legacy includes a flow of microbial genes across generations), (iii) the overall the degree of phylogenetic similarity of their microbiota was not significantly greater among mono- versus dizygotic pairs (emphasizing the importance of early environmental exposures in defining gut microbial community structure), and (iv) despite interpersonal variations in the representation of bacterial species, these different species assemblages share a common set of microbial genes (there is a 'core microbiome'). Intriguingly, Peter's analysis provided the first evidence that bacterial diversity in the gut communities of obese individuals is significantly less than in lean individuals, leading us to hypothesize that 'job vacancies' (unfilled niches) may contribute to obesity and its associated metabolic abnormalities. He also documented significant differences in the representation of genes involved in various facets of nutrient processing/metabolism in lean versus obese microbiomes.

Peter and Vanessa Ridaura, another very bright PhD student in the lab, went on to develop methods for transplanting and reliably replicating previously frozen fecal microbiota from a given human donor in formerly germ-free mice (Turnbaugh et al., 2009b). This was a 'eureka moment' for us: it meant that we could (with appropriate human studies approval) take a picture ('snapshot') of the gut community of subjects representing a given different age, geographic location, cultural tradition and health state, convert this single snap shot to multiple snapshots (through replication of the intact uncultured community in gnotobiotic animals) and then convert the snapshot to a 'movie' by following the recipient mice over time under defined conditions.

A particularly intriguing opportunity became apparent to us at this time: we could not only transplant the donor's microbial community to recipient gnotobiotic mice but also his or her diet, cooked in ways that represented the culinary traditions of their human communities. We reasoned that this would allow us to test 'diet-by-microbiota-by-host' interactions in a number of different contexts, including diets with systematically varied ingredients, to identify relationships between specific diet components and community members. We dreamed about the possibility that this approach could be leveraged to discover/develop affordable health-promoting foods based on consideration of the targeted consumers' microbiota (see below).

Andy Goodman, a very talented post-doc working with others in the lab, developed techniques for culturing the majority of (anaerobic) bacterial taxa from previously frozen human fecal microbiota samples and creating clonally-arrayed 'personal' culture collections (Goodman et al., 2011). These collections were established in multi-well plates where each well contained a single recovered bacterial strain from the donor:

the identity of the strain could be defined by sequencing its phylogenetic marker gene encoding 16S rRNA or its entire genome characterized. Each of these culture collections was composed of organisms that had assembled together in a given human; they represented the evolutionary history of that gut community during the life of that person (Faith et al., 2013). They also reflected the environmental microbial reservoirs, diet and other selective pressures experienced by people living in the donor's geographic and cultural milieu.

Andy showed these personal culture collections could be transplanted into gnotobiotic mice and the effects of diet manipulations on the component organisms defined. He also developed ways to perform genome-wide transposon mutagenesis of selected cultured members of the microbiota. Germ-free mice were colonized with a library of tens of thousands of these isogenic mutants, each with a single transposon insertion, together with other wild-type organisms representing other phylogenetic lineages normally present in the human gut community. Using animals harboring these model communities, Andy was able to identify genes in the organism of interest that conferred fitness in specified diet and community contexts (Goodman et al., 2009).

Armed with these methods, Vanessa Ridaura, a very courageous and dedicated PhD student, found that intact uncultured fecal microbiota samples obtained from twins stably discordant for obesity could transmit the donors' discordant adiposity phenotypes as well as obesity-associated metabolic abnormalities to recipient gnotobiotic mice (Ridaura et al., 2013). She then showed that personal culture collections generated from the donors' fecal microbiota could do the same. Co-housing these gnotobiotic mice shortly after they received their human microbiota transplants resulted in invasion of bacterial species from cagemates with the lean co-twin's community into the guts of cagemates harboring the obese twin's microbiota (and not vice versa). Invasion was associated with prevention of an increased body mass/adiposity phenotype and metabolic abnormalities associated with human obesity. The principal invaders belonged to the Bacteroidetes. Invasion and prevention of the transmitted obesity/metabolic phenotypes were dependent upon diet, occurring when mice were consuming a representative cooked USA diet low in saturated fats and high in fruits and vegetables but not with one high in saturated fats and low in fruits and vegetables (Ridaura et al., 2013).

While the causes of human obesity are multifactorial (caloric intake, the amount of energy expended, host genetics), these studies suggest that in some humans the microbiota is shifted to a reduced diversity state that can sustain obesity and its associated metabolic abnormalities. Filling job vacancies (niches) in their microbiota with

bacterial strains represented in lean individuals requires dietary ingredients that permit these microbes to establish themselves and express their heath-promoting functions. Vanessa's approach offers the promise of facilitating identification of next generation probiotics and synbiotics (combinations of pre- and probiotics). Her results emphasize that new foods that promote health should be designed from the 'inside-out' not just the 'outside in'; i.e., they should take into account the consumer's microbiota and its capacity for transforming food ingredients into metabolic products that beneficially influence host biology. Preclinical proof-of-concept can then be translated to tests in humans, starting with the very population whose microbiota and foods were used to fashion the gnotobiotic animal model.

Jeremiah Faith, a visionary post-doc in the lab with a background in applied mathematics, genomics and microbiology, was able to advance our understanding of how specific diet components affect specific members of the human gut microbiota by using gnotobiotic mice colonized with defined collections of cultured (and sequenced) human bacterial strains and feeding them a sequence of different diets composed of ingredients with systematically varied concentrations. Different mice, all harboring the same defined human gut community, received the different diets in different order, allowing replication of diet exposures across different mice while at the same time testing for hysteresis effects. Applying simple linear models and feature selection, he was able characterize (and predict) which components of a diet can affect the representation of a given community member: a powerful illustration of this approach came from his study involving different human baby foods (Faith et al., 2011).

Jeremiah and Philip Ahern, post-doc with a background in immunology, also developed a scalable, unbiased approach for identifying human gut bacterial strains capable of conveying physiological, metabolic or immunological effects (Faith et al., 2014). Their method involves generating a clonally arrayed collection of sequenced bacterial strains from a human donor whose intact uncultured microbiota transmits phenotypes of interest to gnotobiotic mice. After establishing that phenotypic transmission can occur with the complete culture collection, the collection is randomly fractionated into subsets of varying size (and overlapping membership). Each subset is gavaged into a different germ-free animal (or group of animals). By repeating this combinatorial search across many subsets, the effects of each strain in the collection can be assayed in the context of diverse community memberships and sizes. Feature selection algorithms and follow-up mono-colonizations are used to identify strains whose presence/absence best explains observed phenotypic variation. This approach, which has

yielded bacterial strains that regulate adiposity, plus various facets of metabolism and immunity (Faith et al., 2014), should facilitate discovery of next-generation probiotics and realization of the microbiota's diagnostic potential.

Over the course of the past 20 years, many talented lab members have shown how gnotobiotic mice, and other species such as zebrafish (Rawls et al., 2004, 2006), can be used to link *in vitro* with *in vivo* genetic, functional genomic, proteomic, metabolic and physiologic analyses of microbial-microbial and microbial-host interactions (e.g., Mahowald et al., 2009; Martens et al., 2009a,b, 2011; McNulty et al., 2011, 2013; Rey et al., 2013; Reyes et al., 2013).

During this journey through the gut microbiota, we never lost our interest in the developmental biology of this microbial 'organ'. We wondered whether there was a definable sequence of assembly of bacterial strains in the guts of healthy infants and children. If this developmental program could be described, was it shared across biologically unrelated individuals living in different parts of the world or did it vary based on geography/cultural traditions? With these questions in mind, we expanded our efforts to understand to the role of the developing microbiota in defining nutritional status. Childhood undernutrition is a pervasive global health problem that reflects a complex dynamic of known and yet to be specified intra- and intergenerational factors. While therapeutic food interventions have reduced mortality, incomplete restoration of healthy growth, including persistent stunting, immune dysfunction, and impaired cognitive development, remains a major problem in children with undernutrition worldwide, preventing them from achieving their full human potential.

Tanya Yatsunenko, a PhD student in the lab, began to test the hypothesis that the gut microbiota is a causal factor in childhood undernutrition. Working with a number of colleagues, she used culture-independent metagenomic methods to characterize the gut (fecal) communities of over 500 infants, children, and adults living on three continents and representing very distinct cultural traditions. She found that it takes 2-3 years after birth for the microbiota to mature to an adult-like configuration in each population she surveyed. Moreover, during this period of postnatal development, there were discernible temporal changes (features of which were shared across the populations) in the proportional representation of microbial genes involved in various aspects of nutrient metabolism (Yatsunenko et al., 2012).

Tanya performed a follow-up study with our clinical collaborators, Mark Manary and Indi Trehan, of over 300 twin pairs, aged 0-3 years, living in five rural villages in Malawi. It revealed that discordance rates for moderate or severe undernutrition were remarkably

high but not significantly different for mono- versus dizygotic twins, highlighting the role environmental factors. Her analyses of twin pairs discordant for kwashiorkor, a form of severe acute malnutrition, disclosed that co-twins with kwashiorkor had gut microbiomes whose configuration appeared younger (less mature) than those of their healthy twin siblings. While a commonly used peanut-based ready-to-use therapeutic food (RUTF) promoted evolution towards a more mature state, the effect was not durable, ceasing once RUTF was withdrawn (Smith et al., 2013).

A causal relationship between the gut community and severe acute malnutrition was established by transplanting intact uncultured fecal samples from co-twins in pairs discordant for kwashiorkor into adult germ-free mice and feeding recipient animals a representative nutrient-deficient Malawian diet. The results disclosed that microbiota from the kwashiorkor but not the healthy co-twin transmitted a weight loss phenotype, metabolic derangements, and an enteropathy characterized by disrupted gut mucosal barrier function (Smith et al., 2013; Kau et al., 2015). The transplanted and replicated kwashiorkor microbiota were structurally and metabolically labile, reconfiguring with RUTF but not in a sustained way. Mice harboring a kwashiorkor microbiota fed a nutrient-sufficient diet did not manifest the phenotypes described above, highlighting the importance of diet-by-microbiota interactions.

Looking to a different part of the world where undernutrition is also pervasive, Sathish Subramanian, an innovative MD/PhD student in the lab, together with our colleagues at the International Centre for Diarrhoeal Disease Research in Bangladeshi (ICDDR,B), collected monthly fecal samples from birth through the first two years of life from twin pairs and from ‘singletons’ who exhibited healthy growth (defined using anthropometric parameters developed by the World Health Organization; http://www.who.int/childgrowth/mgrs/en/). Applying culture-independent methods, he identified bacterial strains whose relative abundances characterize normal gut community development. Using machine-learning methods, these strains were incorporated into a sparse Random Forests (RF)-based model containing the most age-discriminatory strains (**Figure 2A,B**). Remarkably, the Bangladeshi-derived model of gut microbial community development (‘maturation’) generalizes to biologically unrelated infants/children with healthy growth phenotypes living in other countries with different cultural traditions (e.g., Malawi; Subramanian et al., 2014).

Sathish used the model to develop two new metrics, a ‘relative microbiota maturity index’ and a ‘microbiota-for-age *Z* (MAZ) score’ that define a child’s gut microbial community development relative to

their chronologic age (**Figure 2C**). He determined that severe acute malnutrition in Bangladeshi children is associated with significant community immaturity that is only transiently and incompletely improved with existing RUTF interventions. He also discovered that microbiota immaturity is also evident in moderate acute malnutrition and correlates with anthropometric measurements (Subramanian et al., 2014).

Laura Blanton, another very gifted PhD student in the lab, working together with Mark Charbonneau, has transplanted fecal samples from Malawian children with varying degrees of growth faltering and microbiota immaturity (collected by our colleagues Mark Manary, Indi Trehan, Per Ashorn, Ken Maleta and Kay Dewey), into young, recently weaned, germ-free mice fed a nutrient-deficient diet representative of that consumed by the microbiota donors. Her work is revealing that age-indicative bacterial strains are also growth indicative.

Together, these studies provide a microbial measure of healthy human postnatal development and suggest that the enduring sequelae/morbidities of undernutrition not repaired by current treatments may reflect their persistent gut community immaturity. They also provide new metrics for classifying undernourished states, potentially before disease becomes severe and gut ecosystem restoration more challenging. The metrics that Sathish developed can be used for assessing the impact of current therapies, as well as new types of food and microbial interventions designed to achieve more durable repair of gut community immaturity and improved clinical outcomes. Gnotobiotic mice harboring immature microbiota from undernourished donors provide preclinical models for directly comparing the effects of several different interventions - something that cannot be done directly in a randomized clinical study where an individual donor is assigned to one or another but not multiple treatment arm(s).

With globalization, we are witnessing dramatic alterations in how and what we eat. Studies of the human gut microbiome will likely have a disruptive effect on our views of human nutrition (Subramanian et al., 2015). Gnotobiotic models of the types described above, incorporating the gut communities of individuals representing different ages, physiologic phenotypes and lifestyles, can provide one route for deriving new definitions of the nutritional value of foods currently being produced for geographically and culturally distinct consumer populations, as well as for the products of discovery and development efforts designed to yield new affordable food sources with improved nutritional value. Results emanating from these preclinical models could inform the design and interpretation of clinical studies designed with the intention of establishing rigorous scientific evidence for any health claims.

The hope is that these efforts will help those involved in agriculture, the food industry and governments consider how to feed a human population, estimated to grow to 9-10 billion by mid-century, in the face of challenges to sustainable agriculture. Considering this nexus between agriculture, food production, food safety, nutritional status and microbiota, it seems timely to suggest that if the microbiota is to be considered as a factor in ongoing plans for developing improved local and global food systems, a holistic and proactive approach is required that includes among other things, consumer education with an evolved vocabulary that meaningfully describes diet-by-microbiota interactions, a regulatory system prepared to process health claims, and integration across many disciplines (including anthropology; Benezra et al., 2011).

Current human microbiome research is addressing questions as old as the discipline of microbiology. It is doing so with new and rapidly expanding sets of experimental and computational tools. We are seeing how it provides an opportunity to forge new alliances across existing disciplines, to spawn new areas of research, to provide new insights about how, with globalization, our rapidly changing cultural traditions and lifestyles are impacting our biology, and to inspire students to expand their understanding of the human condition so that they can devote themselves to addressing some of the most challenging and pressing global health problems we face this century (Xu and Gordon 2003, Gordon, 2012). With so much public expectation (and enthusiasm) for this area of research, including its potential for yielding new strategies for disease treatment or prevention, there is need for sobriety. Much more work needs to be done to (i) understand the mechanisms that mediate our interactions with our indigenous microbes, (ii) provide rigorous proof for causal relationships between a given microbiota configuration or configurations and our human biological features, (iii) ascertain the generalizability of the effects of microbiota-targeted interventions across different populations, and (iv) sponsor thoughtful, timely and open discussions of the ethical, social, safety, regulatory, and commercial issues raised by this research (Gordon, 2012; Subramanian et al., 2015).

Reflecting on our lab's journey and looking forward, I believe that all the work that has been done so far makes it possible for us to just now begin to do what we have always dreamed of doing. The best part of the journey is now and will be tomorrow.

Acknowledgements

I have mentioned only a few of the many people who have shared this journey. I would like to express my deepest appreciation to the 59 students who have done and are doing their thesis work in the lab, and the 63 post-docs who have chosen to spend several years of their lives sharing their insights, dreams, passion for work and discovery with one another and myself. I am grateful to Michael Barratt for his wise stewardship of our international clinical collaborations, Sabrina Wagoner, Jill Manchester for being such wonderful lab managers, Jiye Cheng, Su Deng and Jessica Hoisington-Lopez for their amazing technical skills, Laura Kyro for graphics support, Stephanie Amen for her administrative support, and to many colleagues and collaborators, including Andrew Heath, Andrew Serazin, Phil Needleman, Rob Knight, Chris Newgard, Bernard Henrissat, Tahmeed Ahmed and the team at ICDDR,B, members of the Breast Milk, Gut Microbiome and Immunity Project sponsored by the Bill and Melinda Gates Foundation, and others too numerous to list. This work would not have been possible without the generous support of a number of funding agencies, including the W.M. Keck Foundation, the Bill & Melinda Gates Foundation, the Crohns and Colitis Foundation of America, the NIH (NIDDK) and several industrial partners, or those individuals, living in various parts of the world, who so kindly agreed to provide samples of their gut communities so that we could come to better appreciate what a splendid mixture of microbial and human parts we are, and how our microbial selves can provide a means to achieve improved health.

References

Bäckhed, F., Ding, H., Wang, T., Hooper, L.V., Koh, G. Y., Nagy, A., Semenkovich, C.F. Gordon, J. I. The gut microbiota as an environmental factor that regulates fat storage. ***Proc. Natl. Acad. Sci. USA*** *101:*15718-15723 (2004).

Bäckhed, F., Manchester, J.K., Semenkovich, C.F., and Gordon, J.I. Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. ***Proc. Natl. Acad. Sci. USA*** *104*: 979–984 (2007).

Benezra, A., DeStefano, J., and Gordon, J.I. Anthropology of microbes, ***Proc. Natl. Acad. Sci. USA*** *109*: 6378-6381(2012).

Bry, L., Falk, P.G., Midtvedt, T., and Gordon, J.I. A model of host-microbial interactions in an open mammalian ecosystem. ***Science*** *273*: 1380-1383 (1996).

Faith, J.J., McNulty, N.P., Rey, F.E., and Gordon, J.I. Predicting a human gut microbiota's response to diet in gnotobiotic mice. ***Science*** *333*: 101-104 (2011).

Faith, J.J., Guruge, J.L., Charbonneau, M., Subramanian, S., Seedorf, H., Goodman, A.L., Clemente, J.C., Knight, R., Heath, A.C., Leibel, R.L., Rosenbaum, M., and Gordon, J.I. The long-term stability of the human gut microbiota. ***Science*** *341*: 1237439 doi: 10.1126/science.1237439 (2013).

Faith, J.J., Ahern, P.P., Ridaura, V.K., Cheng, J., and Gordon, J.I. Identifying gut microbiome-host phenotype relationships using combinatorial communities in gnotobiotic mice, ***Science Translational Med*** *6*: 220ra11 (2014).

Goodman, A.L., McNulty, N.P., Zhao, Y., Leip, D., Mitra, R.D., Lozupone, C.A., Knight, R., and Gordon, J.I. Identifying genetic determinants needed to establish a human gut symbiont in its habitat. ***Cell Host & Microbe*** *6*:279-289 (2009).

Goodman, A.L., Kallstrom, G., Faith, J.J., Reyes, A., Moore, A., Dantas, G., and Gordon, J.I. Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. ***Proc. Natl. Acad. Sci. USA***, *108*: 6252-6257 (2011).

Gordon, J.I. Intestinal epithelial cell differentiation: New insights from chimeric and transgenic mice. ***J. Cell Biol.*** *108*: 1187-1194 (1989).

Gordon, J.I., Honor thy gut symbionts redux. ***Science*** *336*: 1251-1253 (2012).

Hermiston, M.L., Green, R.P., and Gordon, J.I. Chimeric-transgenic mice represent a powerful tool for studying how the proliferation and differentiation programs of intestinal epithelial cell lineages are regulated. ***Proc. Natl. Acad. Sci. USA*** *90*: 8866-8870 (1993).

Hermiston, M.L., and Gordon, J.I. *In vivo* analysis of cadherin function in the mouse intestinal epithelium: Essential role in adhesion, maintenance of differentiation, and regulation of programmed cell death. ***J. Cell Biol.*** *129*: 489-506 (1995).

Hermiston, M.L., Wong, M.H., and Gordon, J.I. Forced expression of E-cadherin in the mouse intestinal epithelium slows cell migration and provides evidence for non-autonomous regulation of cell fate in a self-renewing system. ***Genes & Dev.*** *10*: 985-996 (1996).

Hooper, L.V., Xu, J., Falk, P.G., Midtvedt, T., and Gordon, J.I. A molecular sensor that allows a gut commensal to control its nutrient foundation in a competitive ecosystem. ***Proc. Natl. Acad. Sci. USA***, *96*: 9833-9838 (1999).

Hooper, L.V., Wong, M.Hs., Thelin, A., Hansson, L., Falk, P.G, and Gordon, J.I. Molecular analysis of commensal host-microbial relationships in the intestine. ***Science*** *291*: 881-884 (2001a)
Hooper, L.V. and Gordon, J.I. Commensal host-bacterial relationships in the gut. ***Science*** *292*:1115-1118 (2001b).

Kau, A.L., Planer, J.D., Liu, J., Rao, S., Yatsunenko, T., Trehan, I., Manary, M.J., Liu, T-C., Stappenbeck, T.S., Maleta, K.M., Ashorn, P., Dewey, K.G., Houpt, E.R., Hsieh, C-S., and Gordon, J.I. Functional characterization of IgA-targeted bacterial taxa from undernourished Malawian children that produce diet-dependent enteropathy. ***Science Translational Med***. (2015).

Ley, R.E., Bäckhed, F., Turnbaugh, P., Lozupone, C.A., Knight, R.D., and Gordon, J.I. Obesity alters gut microbial ecology. ***Proc. Natl. Acad. Sci. USA*** *102*: 11070-11075 (2005).

Ley, R.E., Turnbaugh, P.J., Klein, S., and Gordon, J.I. Microbial ecology: human gut microbes associated with obesity. ***Nature*** *444*: 1022-1023 (2006).

Ley, R.E., Hamady, M., Lozupone, C., Turnbaugh, P.J., Ramey, R.R., Bircher, J.S., Schlegel, M.L., Tucker, T.A., Schrenzel, M.D., Knight, R., and Gordon, J.I. Evolution of mammals and their gut microbes. ***Science*** *320*:1647-1651 (2008).

Mahowald, M.A., Rey, F.E., Seedorf, H., Turnbaugh, P.J., Fulton, R.S., Wollam, A., Shah, N., Wang, C., Magrini, V., Wilson, R.K., Cantarel,

B.L. Coutinho, P.M., Henrissat, B., Crock, L.W., Russell, A., Verberkmoes, N.C., Hettich, R.L., and Gordon, J,I., Characterizing a model human gut microbiota composed of members of its two dominant bacterial phyla. ***Proc. Natl. Acad. Sci. USA*** *106*: 5859-5864 (2009).

Martens, E.C., Chiang, H.C., and Gordon, J.I. Mucosal glycan foraging enhances fitness and transmission of a saccharolytic human gut bacterial symbiont. ***Cell Host & Microbe*** *4*: 447-457 (2008).

Martens, E.C., Roth, R., Heuser, J., and Gordon, J.I. Coordinate regulation of glycan degradation and capsule synthesis by a prominent human gut symbiont. ***J. Biol. Chem***. *248*: 18445-57 (2009).

Martens, E.C., Lowe, E., Abbott, W., Chiang, H., Pudlo, N., Wu, M., McNulty, N., Abbott, W., Henrissat, B., Gilbert, H.J., Bolam, D.N., and Gordon, J.I. Recognition and degradation of plant cell wall polysaccharides by two human gut symbionts. ***PLoS Biol.*** *9*: e10001221 (2011).

McNulty, N, Yatsunenko, T, Hsiao, A., Faith, J., Muegge, B., Goodman, A., Henrissat, B., Oozeer, R., Cools-Portier, S., Gobert, G., Chervaux, C., Knights, D., Lozupone, C., Knight, R., Duncan, A.E., Bain, J.R., Muehlbauer, M.J., Newgard, C.B., Heath, A.C., and Gordon, J.I. The impact of a consortium of fermented milk strains on the gut microbiome of gnotobiotic mice and monozygotic twins. ***Science Translational Med.*** *3*: 106ra106 (2011).

McNulty, N.P., Wu, M., Erickson, A.R., Martens, E.C., Pudlo, N.A., Muegge, B., Henrissat, B., Hettich, R.L., and Gordon, J.I. Effects of diet on resource utilization by a defined model human gut microbiota containing *Bacteroides cellulosilyticus* WH2, a symbiont with an extensive glycobiome, ***PLoS Biol.*** *11*: e1001637 (2013).

Rawls, J.F., Samuel, B.S., and Gordon, J.I. Gnotobiotic zebrafish reveal evolutionarily conserved responses to the gut microbiota. ***Proc. Natl. Acad. Sci. USA*** *101*: 4596-4601 (2004).

Rawls, J.F., Mahowald, M.A., Ley, R.E., and Gordon, J.I. Reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients reveal host habitat selection. ***Cell*** *127*: 423-433 (2006).

Rey, F.E., Gonzalez, M., Cheng, J., Wu, M., Ahern, P.P., and Gordon, J.I. The metabolic niche of a prominent sulfate-reducing human gut bacterium. ***Proc. Natl. Acad. Sci. USA*** *110*: 13582-13587 (2013).

Reyes, A., Wu, M., McNulty, N.P., Rohwer, F., and Gordon, J.I. Gnotobiotic mouse model of phage-bacterial host dynamics in the human gut. ***Proc. Natl. Acad. Sci. USA*** *110*: 20236-20241 (2013).

Ridaura, V.K., Faith, J.J., Rey, F.E., Cheng, J., Duncan, A.E., Kau, A.L., Lombard, V., Henrissat, B., Bain, J.R., Muehlbauer, M.J., Ilkayeva, O., Ursell, L.K., Clemente, J.C., Van Treuren, W., Walters, W.A., Newgard, C.B., Knight, R., Heath, A.C., and Gordon, J.I. Cultured gut microbiota from twins discordant for obesity modulate adiposity and metabolic phenotypes in mice. ***Science*** *341*: 1241214 (2013)

Sonnenburg, J.L., Xu, J., Leip, D.D., Chen, C.H., Westover, B.P., Weatherford, J., Buhler, J.D., and Gordon, J.I. Glycan foraging *in vivo* by an intestine-adapted bacterial symbiont. ***Science*** *307*: 1955-1959 (2005).

Smith, M.I., Yatsunenko, T., Manary, M.J., Trehan, I., Mkakosya, R., Cheng, J., Kau, A., Rich, S.S., Concannon, P., Mychaleckyj, J.C., Liu, J., Houpt, E., Li, J.V., Holmes, E., Nicholson, J., Knights, D., Ursell, L.K., Knight, R., and Gordon, J.I. Gut microbiomes of Malawian twin pairs discordant for kwashiorkor. ***Science*** *339*: 548-554 (2013).

Subramanian, S., Yatsunenko, T., Huq, S., Haque, R., Mahfuz, M., Alam, M.A., Benezra, A., DeStefano, J., Meier, M.F., Muegge, B.D., Barratt, M.J., Zhang, Q., Province, M.A., Petri, W.A., Ahmed, T., and Gordon, J.I. Persistent gut microbiota immaturity in malnourished Bangladeshi children. ***Nature*** *509*: 417-421 (2014).

Subramanian, S., Blanton, L., Frese, S.A., Charbonneau, M., Mills, D.A., and Gordon, J.I. Cultivating healthy growth and nutrition through the gut microbiota. ***Cell*** (2015).

Sweetser, D.A., Birkenmeier, E.H., Hoppe, P.C., McKeel, D.W., and Gordon, J.I. Mechanisms underlying generation of gradients in gene expression within the intestine: An analysis using transgenic mice containing fatty acid binding protein-human growth hormone fusion genes. ***Genes & Dev.*** *2*: 1318-1332 (1988).

Turnbaugh, P.J.Ley, R.E., Mahowald, M.A., Magrini, V., Mardis, E.R. and Gordon, J.I. An obesity-associated gut microbiome with increased capacity for energy harvest. ***Nature*** *444*:1027-1031 (2006).
Turnbaugh, P.J., Bäckhed, F., Fulton, L. and Gordon, J.I. Diet-induced obesity is linked to marked but reversible alterations in the mouse distal gut microbiome ***Cell Host & Microbe*** *3*: 213-223 (2008).

Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B., Duncan, A., Ley, R.E., Sogin, M.L., Jones, J., Roe, B.A., Affourtit, J.P., Egholm, M., Henrissat, B., Heath, A.C., Knight, R. and Gordon J.I. A core gut microbiome in obese and lean twins. ***Nature*** *457*:480-484 (2009a)
Turnbaugh, P.J., Ridaura, V.K., Faith, J.J., Rey, F.E., Knight, R. and Gordon, J.I. The effect of diet on the human gut microbiome: a metagenomic analysis in humanized notobiotic mice. ***Science Trans. Med.*** *1*: 6ra14 (2009b).

Xu, J., Bjursell, M.K., Himrod, J., Deng, S. Carmichael, L.K., Chiang, H.C., Hooper, L.V., and Gordon, J.I. A genomic view of the human-*Bacteroides thetaiotaomicron* symbiosis. ***Science*** *299*: 2074-2076 (2003).

Xu, J., and Gordon, J.I. Honor thy symbionts. ***Proc. Natl. Acad. Sci. USA*** *100*: 10452-10459 (2003).

Yatsunenko, T., Rey, F.E., Manary, M.J., Trehan, I., Dominguez-Bello, M.G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R.N., Anokhin, A.P., Heath, A.C., Warner, B., Reeder, J., Kuczynski, J., Caporaso, J.G., Lozupone, C.A., Lauber, C., Clemente, J.C., Knights, D., Knight, R., and Gordon, J.I. Human gut microbiome viewed across age and geography. ***Nature*** *486*: 222-227 (2012).

Figure Legends

Figure 1 – The saccharolytic human gut symbiont, *Bacteroides thetaiotaomicron,* is sufficient to efficiently induce a naturally occuring postnatal developmental program of fucosylated glycan production in the distal gut epithelium of gnotobiotic mice if it is capable of utilizing the L-fucose it liberates. (A-H) Wholemount preparations of the distal small intestines of germ-free and conventionally raised mice belonging to the NMRI inbred strain sacrificed at postnatal days 17 (P17), P21, P23, P25, and P28. In the absence of a gut microbiota, production of□□1,2 linked fucosylated glycans in the distal small intestinal (ileal) epithelium [detected as brown

staining cells with the peroxidase conjugated lectin Ulex Europaeus Agglutinin (UEA-1)] does not progress. (**I,J**) Mono-colonization of germ-free NMRI mice with wild-type *B. thetaiotaomicron*, or an isogenic strain with a transposon mutant that renders it unable to utilize L-fucose as a carbon source. Seven days after colonization at P28, the wild-type strain is able to induce the normal developmental program of distal small intestinal fucosylated glycan production. The mutant strain is much less efficient at inducing fucosylation. (Taken from Bry et al., 1996).

Figure 2 – Defining maturation of the human gut microbiota in infants and children (**A**) Summary of approach. Monthly fecal samples were collected from healthy infants and children living in an urban slum in Dhaka, Bangladesh during their first two years of postnatal life. Bacterial strains that discriminate different stages of development were identified by a machine learning-based (Random Forests) regression of bacterial 16S rRNA datasets, generated from these fecal samples, to the respective chronologic ages of the microbiota donors at the time of sample collection. (**B**) The most age-discriminatory taxa, as defined by their feature importance, were used to generate a sparse 24-taxon model whose output ('microbiota age') is a microbiota-based prediction of the chronologic age of a healthy child. The graphs show microbiota age plotted against the chronologic age of healthy children used as a training set to fit the regression (each dot is a fecal sample from an individual child), plus validation sets composed of different groups of children with healthy growth profiles living in the same location that were not used to train the Random Forests model (singletons and a group of twins and triplets). (**C**) Two metrics of microbiota maturation are computed. 'Relative microbiota maturity' is the deviation, in months, from a smooth-spline fit of microbiota age values with respect to chronologic age, fitted using the validation datasets (black dashed curve). The red circle represents a fecal sample, collected from an undernourished child, that is 11 months below the spline fit, indicating negative relative microbiota maturity (i.e., an immature microbiota). A microbiota-for-age *Z* score (MAZ) is computed by dividing the difference between a child's microbiota age and the median microbiota age of healthy controls in the same monthly chronologic age bin over the standard deviation within the same age bin. The median and standard deviation of each bin are computed using the validation datasets. Panels B and C are from Subramanian et al. (2014).

Fig. 1

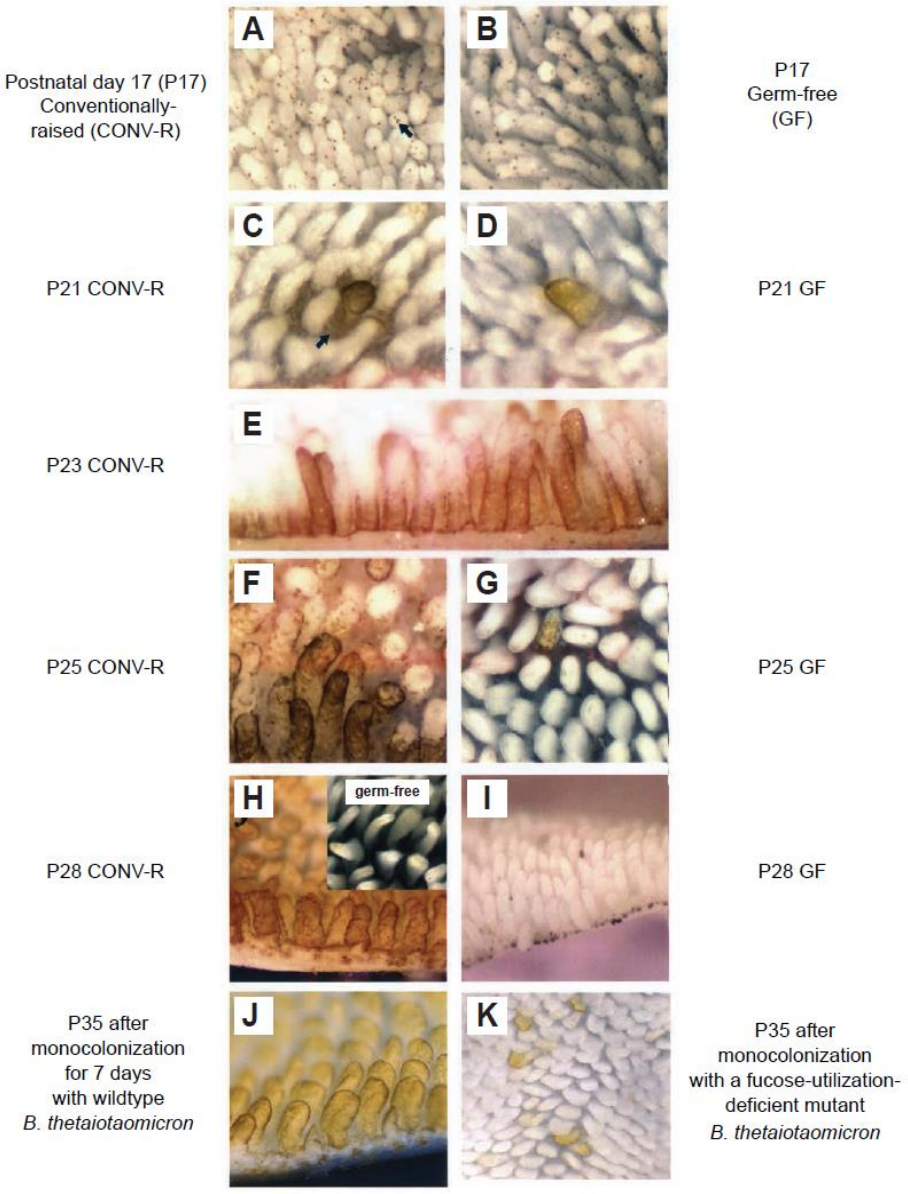
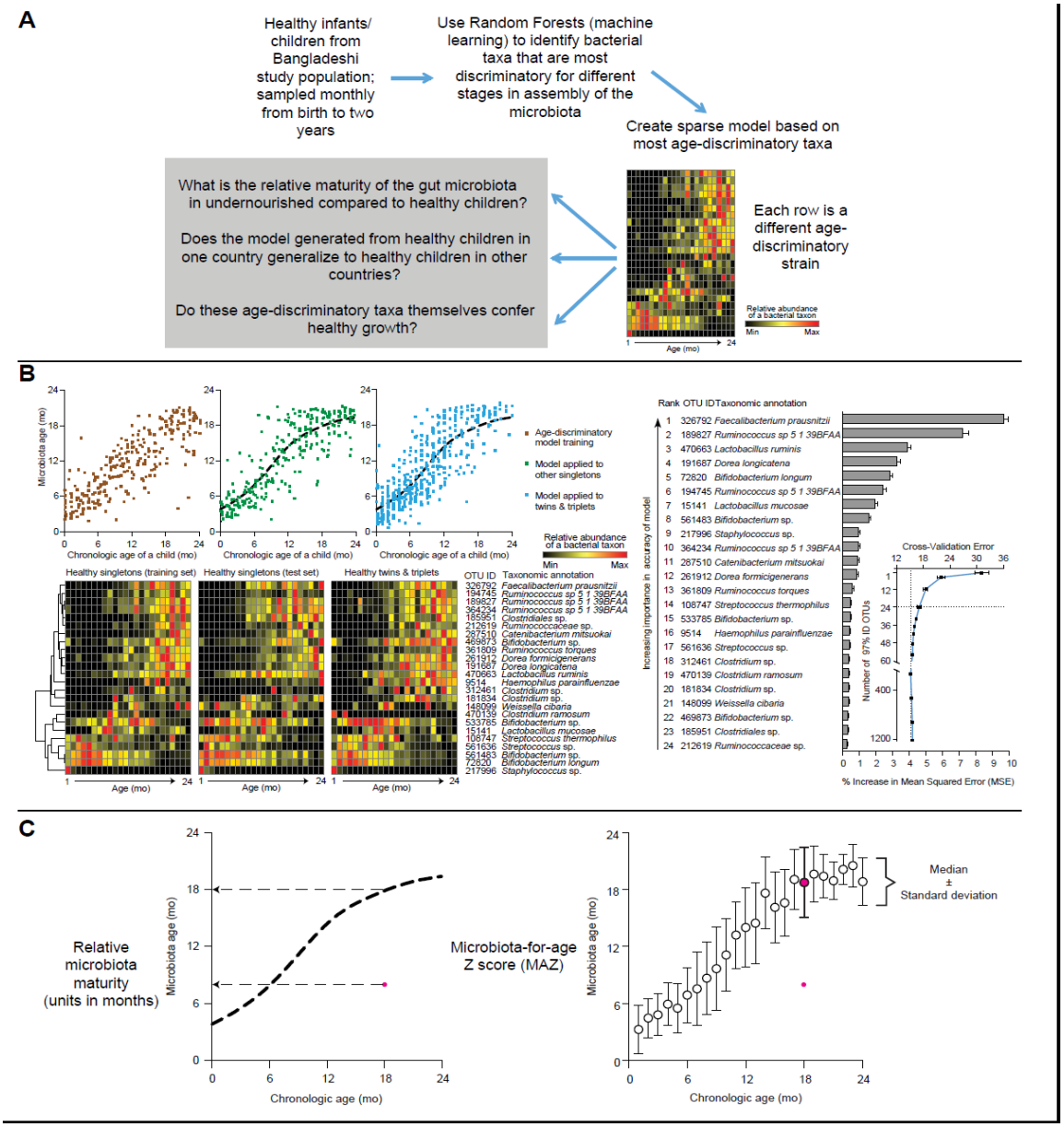
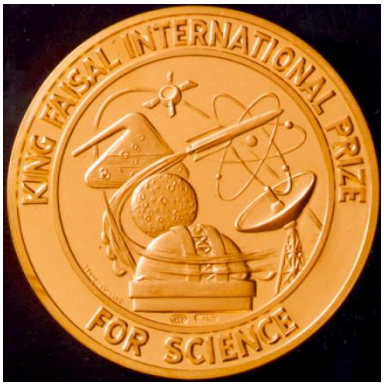

Postnatal day 17 (P17) Conventionally-raised (CONV-R)
P17 Germ-free (GF)
A
B
P21 CONV-R
P21 GF
C
D
P23 CONV-R
E
P25 CONV-R
P25 GF
F
G
P28 CONV-R
P28 GF
H
germ-free
I
P35 after monocolonization for 7 days with wildtype B. thetaiotaomicron
P35 after monocolonization with a fucose-utilization-deficient mutant B. thetaiotaomicron
J
K

Fig. 2


A
Healthy infants/ children from Bangladeshi study population; sampled monthly from birth to two years
Use Random Forests (machine learning) to identify bacterial taxa that are most discriminatory for different stages in assembly of the microbiota
Create sparse model based on most age-discriminatory taxa
What is the relative maturity of the gut microbiota in undernourished compared to healthy children?
Does the model generated from healthy children in one country generalize to healthy children in other countries?
Do these age-discriminatory taxa themselves confer healthy growth?
Each row is a different age-discriminatory strain
Relative abundance of a bacterial taxon
Min
Max
Age (mo)
B
Microbiota age (mo)
Chronologic age of a child (mo)
Age-discriminatory model training
Model applied to other singletons
Model applied to twins & triplets
Relative abundance of a bacterial taxon
Min
Max
Healthy singletons (training set)
Healthy singletons (test set)
Healthy twins & triplets
OTU ID
Taxonomic annotation
326792 Faecalibacterium prausnitzii
194745 Ruminococcus sp 5 1 39BFAA
189827 Ruminococcus sp 5 1 39BFAA
364234 Ruminococcus sp 5 1 39BFAA
185951 Clostridiales sp.
212619 Ruminococcaceae sp.
287510 Catenibacterium mitsuokai
469873 Bifidobacterium sp.
361809 Ruminococcus torques
261912 Dorea formicigenerans
191667 Dorea longicatena
470663 Lactobacillus ruminis
9514 Haemophilus parainfluenzae
312461 Clostridium sp.
181834 Clostridium sp.
148099 Weissella cibaria
470139 Clostridium ramosum
533785 Bifidobacterium sp.
15141 Lactobacillus mucosae
108747 Streptococcus thermophilus
561636 Streptococcus sp.
561483 Bifidobacterium sp.
72820 Bifidobacterium longum
217996 Staphylococcus sp.
Age (mo)
Rank OTU ID Taxonomic annotation
1 326792 Faecalibacterium prausnitzii
2 189827 Ruminococcus sp 5 1 39BFAA
3 470663 Lactobacillus ruminis
4 191687 Dorea longicatena
5 72820 Bifidobacterium longum
6 194745 Ruminococcus sp 5 1 39BFAA
7 15141 Lactobacillus mucosae
8 561483 Bifidobacterium sp.
9 217996 Staphylococcus sp.
10 364234 Ruminococcus sp 5 1 39BFAA
11 287510 Catenibacterium mitsuokai
12 261912 Dorea formicigenerans
13 361809 Ruminococcus torques
14 108747 Streptococcus thermophilus
15 533785 Bifidobacterium sp.
16 9514 Haemophilus parainfluenzae
17 561636 Streptococcus sp.
18 312461 Clostridium sp.
19 470139 Clostridium ramosum
20 181834 Clostridium sp.
21 148099 Weissella cibaria
22 469873 Bifidobacterium sp.
23 185951 Clostridiales sp.
24 212619 Ruminococcaceae sp.
Increasing importance in accuracy of model
Cross-Validation Error
Number of 97% ID OTUs
% Increase in Mean Squared Error (MSE)
C
Relative microbiota maturity (units in months)
Microbiota age (mo)
Chronologic age (mo)
Microbiota-for-age Z score (MAZ)
Median ± Standard deviation

WINNERS OF THE 2015
KING FAISAL INTERNATIONAL PRIZE
FOR SCIENCE



The King Faisal International Prize for Science for this year 2014G (1435H), Topic: Mathematics, has been awarded to: Professor Gerd Faltings (Germany) Director at the Max-Planck Institute for Mathematics in Bonn.

Professor Faltings has made groundbreaking contributions to algebraic geometry and number theory. His work combines ingenuity, vision and technical power. He has introduced stunning new tools and techniques which are now constantly used in modern mathematics. His deep insights into the p-adic cohomology of algebraic varieties have been crucial to modern developments in number theory. His work on moduli spaces of abelian varieties has had great influence on arithmetic algebraic geometry. He has introduced new geometric ideas and techniques in the theory of Diophantine approximation, leading to his proof of Lang's conjecture on rational points of abelian varieties and to a far-reaching generalization of the subspace theorem. Professor Faltings has also made important contributions to the theory of vector bundles on algebraic curves with his proof of the Verlinde formula

Diophantine Equations and Beyond

Professor Gerd Faltings
Max Planck Institute for Mathematics
Vivatsgasse 7
D-53111 Bonn
Germany

1. Introduction

This article gives an overview of my research. The most important results are io diophantine equations, but they led to other fields and I trie to explain how this happened.

2. Diophantine equations

Diophantine problems deal with solutions of algebraic equations in rational num- bers. Recall that the natural numbers {1, 2, 3...} are obtained by simple counting, and they suffice for some purposes. However for applications

one usually has to con- struct more complicated numbers: First one needs the zero and negative numbers
$\{0, -1, ...\}$, then rational numbers a/b where a and b are integers with b different from 0. The real numbers R are obtained as limits of rational numbers, as for
example the squareroot $\sqrt{2}$ or the number π. Finally for complex numbers C one
has to add a squareroot i of -1, that is they are linear combinations $a + bi$. with a and b real numbers. For many purposes the complex numbers suffice. For example any algebraic equation has a root in C. However sometimes one has to replace the real numbers by different completions of the rationals. For any prime p consider two rationals a/b and a'/b as p-adically close if the integer $a - a'$ is divisible by a big power of p. Then adding p-adic limits extends the rationals Q to the field Q_p of p-adic numbers. For example the series

$$1 + 2 - 2 + 24 - , , , = 1 - \sum (-2)^n (1 \cdot 3 \cdot (2n-1))/(1 \cdot ...n) \; n \geq 1$$

has a limit in Q_2 which is equal to $\sqrt{5}$. Q_p has some similarity to the reals R but is also quite different in certain as-
pects. For example it does not suffice to add one element (like i) to it to make it
algebraically closed. Instead one has to adjoin infinitely many elements and add a further completion to extend Q_p to a complete algebraically closed field C_p.
Diophantine geometry deals with solutions of algebraic equations in integers
or rational numbers. Examples are Pythagorean triples (solutions in integers of $a^2 + b^2 = c^2$, or in rationals of $x^2 + y^2 = 1$), or the Fermat equation ($a^n + b^n = c^n$, or $x^n + y^n = 1$, $n \geq 3$). In general an algebraic variety is the set of common solutions of finitely many polynomial equations. Examples are $x^2 - y^2 = 1$ or $x^2 + y^3 + z^5 = 0$, but not $3^x - 2^y = 1$ (it involves nonalgebraic functions). The algebraic variety is defined over Q if the polynomials defining it have coefficient in Q. It then has points (that is common solutions of the equations) in any overfield of Q, that there are Q-rational, R-rational and C-rational points. For example the algebraic variety given by $x^2 + y^2 = -1$ is defined over Q but has no rational points over Q or even over R. On the other hand over C it is isomorphic to $x^2 + y^2 = 1$

(multiply x and y by i) and has many C-rational points. This illustrates that over the complex numbers C many things become simpler. An algebraic variety is called smooth if the complex points form a manifold. For example the variety given by $x^2 + y^2 = 1$ is smooth while $y^2 - x^3 = 0$ defines a nonsmooth variety.

An important invariant of an algebraic variety is its dimension. It roughly says on how many complex parameters the C-rational points depend. Varieties of dimension zero are finite sets. In the next case of dimension one the varieties are called curves. Such a curve has an important invariant, the genus. Rational points on curves of genus zero can be parametrised. For example rational solutions to $x^2 + y^2 = 1$ are of the form $x = (1 - t^2)/(1 + t^2)$, $y = 2t/(1 + t^2)$. For curves of genus one Mordell showed that rational points form a finitely generated abelian group. That is they are much rarer than for genus zero, but there still may be infinitely many of them. He conjectured that for genus bigger than one the set of rational points is finite.

Important concepts in the study of diophantine equations are height and the notion of good reduction. The hight of a rational number a/b is the maximum of the sizes of the numerator a and the denominator b (assumed to be coprime). The height of an n-tuple $(x_1, ..., x_n)$ of rational numbers is the maximum of the height of the coordinates x_i. The height is important for proving finiteness theorems because it suffices to give an upper bound for the height of solutions. Of course this tends to be difficult for interesting problems.

Good reduction at a prime p means in the simplest case that for a rational number a/b p does not divide the denominator. This can fail only for finitely many primes which are called the primes of bad reduction. Another example: An equation defines a smooth algebraic variety if some discriminant is nonzero. If the equation has integers as coefficients this discriminant is also an integer, and thus only divisible by finitely many primes. These are the primes of bad reduction for this property.

In more generality one may enlarge the rationals by adding solutions of algebraic equations to obtain algebraic numberfields. An example is the field $Q(\sqrt{2})$ which consists of all linear combinations $a+b\sqrt{2}$ with rationals a, b. Note that this field has a symmetry by sending $a + b\sqrt{2}$ to $a - b\sqrt{2}$, quite similar to complex conjugation on C. So if an equation has coefficients in Q its solutions in $Q(\sqrt{2})$ admit this
symmetry. A version of it exists for all numberfields and is called the Galois-group. One of the most powerful tools in diophantine geometry

consists in passing from rational points to representations of the Galois-group.

As an example we illustrate this in the case of the Fermat equation $a^n + b^n = c^n$:

We want to show that it has no nontrivial solutions in integers if $n \geq 3$. It suffices to consider the cases where $n = 4$ or where n is an odd prime. Then we consider the auxiliary Frey elliptic curve, that is solutions of the equation

$$y^2 = x(x - a^n)(x + b^n).$$

If we add a point at infinity the (say) complex solution (x, y complex numbers) form a commutative group, that is we can define an addition of two such points. The group law is determined by the fact that the three intersection points of the curve with any straight line (in the (x, y)-plane) add up to zero. The height of such an elliptic curve is given by the size of the discriminant which is $4a^n b^n c^n$, and the primes of bad reduction are the primes which divide a, b or c. The n-division points (adding the point n-times to itself gives zero) form a subgroup of order n^2. As the group law is given by algebraic equation with coefficients in Q these division points lie in algebraic numberfields and thus admit an action of the Galois-group. This Galois-action has certain special properties:

Namely at the primes of bad reduction one would expect that the representation notices them (it "ramifies"). In our case (the Frey-curve for a Fermat equation) the primes of bad reduction are those dividing one of the numbers a, b, c. However because they occur as n-th powers these primes of bad reduction are not seen by the n-torsion points. In Wiles solution of the Fermat problem the key step is to show that this implies that the elliptic curve is "congruent modulo n" to an elliptic curve with no primes of bad reduction. Actually one has to work with a generalisation of elliptic curves (modular forms), but nevertheless one shows that no such object without primes of bad reduction exists, and thus the Fermat equation has no nontrivial solution.

For more complicated diophantine equations one sometimes can still associated to solutions a Frey elliptic curve. However this construction would help only if one could bound the height of this curve. To achieve this one has to solve the "abc-conjecture", one of the most important open problems in the theory.

For more general diophantine equations we cannot hope for no solutions, but only for qualitative statements like finiteness of solutions, for special types of alge- braic varieties. To achieve this for curves (that is to show the Mordell conjecture) the method of Parshin-Arakelov associated to each

solution a new curve and its Ja- cobian, which is a generalisation of the elliptic curves which we encountered for the Fermat problem. They and Szpiro could show the Mordell conjecture over function fields, which are similar to numberfields but where additional tools are available. Also Szpiro emphasized the importance of Arakelov theory which allows to carry over some techniques from function to numberfields. However one important tool (Kodaira-Spencer classes) was missing.

Quite unexpectedly this difficulty was resolved. namely associated to the auxil- iary curves are their Jacobians and the Galois action on the torsion points of the Jacobians. This Galois action again has primes of bad reduction, and this set of primes is not empty but al least predetermined by the curve and not by the ratio- nal point on it. By the general theory (the Weil conjectures) there are only finitely many such representations. If we show that only finitely many points can give rise to the same representation we derive finiteness of rational points.

Now if two points give rise to the same representation the corresponding Jaco- bians are not the same but at least they are similar. The technical term is that they are isogenous. I could show in 1983 that in a given isogeny class the height of the Jacobian (a measure for its complexity) is bounded, and derives that there can be only finitely many Jacobians in this isogeny class. Thus there are only finitely many rational points.

The proof of the Mordell conjecture led me to further work in two fields where the necessary results could be obtained in an adhoc manner but where a fully satisfactory treatment required further work. One was the need to define heights for abelian varieties. These correspond to rational points on a certain moduli space, but to get a good theory one has to compactify that space, that is to add certain degenerate abelian varieties. Over the complex numbers compactifications had been defined by Baily-Borel, Shimura, and Mumford, but they had no arithmetic interpretation, nor did the construction allow to bound the primes of bad reduction. However Mumford also had found a construction of degenerate abelian varieties,

and I could show that Mumford's construction gives all such degenerations, and allows to define local coordinates at the boundary of the compactifications. The resulting arithmetic compactification is called the toroidal compactificationy. I wrote a book about that in collaboration with C.L.Chai.

Another development was the local theory of p-adic Galois representations. What was necessary for Mordell had already been done by Tate in the one dimensional case. It turned out that his method generalises to higher dimensions and yields a "p-adic Hodge

decomposition". Moreover the method allowed to attack the comparison between etale and crystalline cohomology which had been conjectured by J.M.Fontaine. The whole theory now has been put on a conceptual basis by P.Scholze.

Another (in fact historically earlier) approach to diophantine equtions has been via diophantine approximation. We illustrate that for the example of Roth's theo- rem: If α is an algebraic number (that is it satisfies a polynomial equation) then α cannot be too well approximated by a rational numbers. Namely for any exponent $r > 2$ there are only finitely many rationals a/b with

$$|\alpha - a/b| \leq 1/b^r.$$

For $r = 2$ the theory of continuous fractions gives (for real irrational numbers) infinitely many solutions. For the proof of Roth's theorem one assumes that the assertion is wrong and derives a contradiction. If the assertion is wrong there are infinitely many fractions a/b satisfying the inequality. Among them the denomi- nators b can become arbitrarily large, so one can find e sequence a_i/b_i with the b_i rapidly increasing (which can be made precise but this requires some technicalities). Then one constructs an auxiliary polynomial F in variables $T_1, ..., T_s$ of multide- gree $(d_1, ..., d_s)$ (with d_i approximately proportional to the invers of $\log(b_i)$) which vanishes to high order at $(\alpha, ..., \alpha)$. Finally one derives a contradiction if s is big enough (depending on how close r is to 2) and if the d_i increase rapidly enough.

Thue was the first to apply this method to diophantine geometry. He was fol-
lowed by Siegel. However only Vojta succeeded much later to give a proof for the Mordell conjecture along these lines. Trying to understand his proof I developed a geometric approach which allowed to generalise it to higher dimensional varieties.

References

[1] G.Faltings, Endlichkeitssatze fu̇r abelsche Varietaten u̇ber Zahlk̆orpern, Invent. math. 73 (1983), 349-366

[2] G.Faltings, Almost ́etale extensions, Ast́erisque 279 (2002), 185-270

[3] G.Faltings, Diophantine approximation on abelian varieties, Ann.of Math. 133 (1991), 549-576 [4] G.Faltings, Coverings of p-adic period domains,

[5] G.Faltings, C.L.Chai, Degeneration of abelian varieties,
Springer Verlag, Berlin 1990 J.reine angew. Math. 643 (2010), 111-139

[6] J.M.Fontaine, Modules galoisiennes, modules filtrés et anneaux de Barsotti-Tate, Astérisque
65 (1979), 3-80
[7] P.Scholze, Perfectoid spaces, Publ. math. IHES 116 (2012), 245-313
[8] P.Vojta, Siegel's theorem in the compact case, Annals of Math. 133, 509-548

Max-Planck-Institut für Mathematik, Vivatsgasse 7,
53111 Bonn, Germany