Undergraduate Student, Graduate Student, and Postdoctoral Training Plans:

The ***Doctoral student*** in the Biomedical Sciences Program will work with Dr. Britton and Dr. Greathouse to learn how to setup, run, and maintain the MBRA system. For the first two weeks of the project, Dr. Greathouse and the doctoral student will work with Dr. Britton at his lab at Baylor College of Medicine to learn how to setup and use the MBRA system. This training will include education on each apparatus contained within the MBRA, followed by observation of staff/students in their use of the MBRA to conduct experiments, and finally conducting test experiments on our own with supervision. Following this training, Dr. Greathouse and the doctoral student will setup the MBRA system at Baylor University. Dr. Greathouse will train the doctoral student in microbial DNA extraction, sequencing, bioinformatic analysis, and statistics. Dr. Britton will provide training in using the MBRA system and conducting bacterial invasion experiments. During the second and third years, the doctoral student will work part time with Dr. Lavado to learn how to perform fatty acid extraction and analysis. Further, all mentors will provide training in grant writing, manuscript preparation, presentation skills, and opportunities to travel to present at national conferences. Dr. Greathouse, Dr. Lavado, and Dr. Britton will continue to co-mentor the BMS doctoral student throughout his/her research until completion of their degree.

The ***Postdoctoral fellow*** will work with Dr. Greathouse and Dr. Chia to learn how to use the MBRA system in a similar manner to the training of the doctoral student. For the first year, the postdoctoral fellow will utilize the microbial sequence data for bioinformatic analysis. They will also work with Dr. Lavado to integrate microbial metabolite data with the sequencing data. In the last 2 years of this research project, the postdoctoral fellow will work with Dr. Chia to integrate the microbial and metabolite data to construct an interaction landscape from response to dietary prebiotics. Next, they will use the interaction landscape to identify which prebiotic fibers can support the microbes capable of evicting or preventing *C. difficile* or ETBF engraftment. Dr. Greathouse and Dr. Chia will also provide training in grant writing, manuscript preparation, presentation skills, and opportunities to travel to present at national conferences.

Impact on transformational undergraduate education

Background

Beginning in 2009, the Department of Biology began developing course-based undergraduate research experiences (CURES) to transform the ways in which our students learn. In line with transformational learning, we believe that the world needs a new generation of trained biologists to solve the "Grand Challenges" of our world. These challenges include understanding the organism's role in the environment, assembling the "Tree of Life", developing tools and models to understand the dynamic interconnected systems of life, understanding genomes and how they produce organisms, and understanding the paradox of stability and change in evolutionary history. These challenges are applicable for global health, sustainable agriculture, clean air, water, and energy. These initiatives have continued to grow and currently include approximate 25% of our introductory biology lab sections and several upper-level courses. **The MBRA will**

enrich these courses by allowing students to design creative and relevant experiments and to analyze large genomic data sets.

Summary of Research Topics

1. BIO 1406: Analysis of soil microbial communities for types and concentrations of bacteriophage or specific host organisms related to various locations or conditions.
2. BIO 1106: Analysis of soil microbial communities and experimental design using additions of different types of ciliates, microbes, or pollutant.
3. Independent projects: This technology will provide the opportunity to have large data sets generated in several research labs. Current examples include water samples and sewage effluent samples. The data could also be generated in labs studying cancer or the human microbiome. This technology is a powerful tool that students can use to analyze many different types of samples. The end result of these experiments is a DNA sample that needs bioinformatic analysis. The BIO 1105/1106 curriculum is primed to include this large-scale research analysis, as well as the undergraduate researchers that are mentored in the Department of Biology each year.

Description

1. There are at least 2 research questions that we would address in the undergraduate lab courses in the Department of Biology with this MBRA technology. **First, students in the SEA-PHAGES program (a program first initiated through a Science Education Alliance grant from Howard Hughes Medical Center) could use a soil lysate to not only search for individual bacteriophages, but with the use of the MBRA, students could ask sophisticated questions about what happens to the bacterial community with the addition of certain phages or when the environmental parameters change (e.g. pH or temperature). This program has achieved success in training undergraduates in microbiology and bioinformatics.** The interesting questions concerning how the phage population changes over time and under different conditions could also be addressed. Community interactions are important in the soil, and known to be important in soil health and plant growth. Baylor undergraduates have archived 168 Arthrobacter phage samples and published 15 phage genomes. **This research has generated dozens of publications within the SEA-PHAGES community, many with undergraduate authors.** Having access to this bioreactor would allow for further experiments and increase the research experience and productivity of our undergraduate students. Phage genomic research holds promise for new innovations in human health and ecology. From the early days, phage biology has led the way to discovery and advances in molecular biology and modern biotechnology. For example, the study of just one type of phage (phage lambda), led to discoveries of DNA binding proteins, basic mechanisms of gene regulation and regulatory networks, the

mechanism of site-specific recombination, and much more. Early genetic engineering used phage lambda vectors to clone DNA fragments and many new technologies followed. Phage are the most abundant biological entities on Earth and their diversity is bound to reveal many more discoveries. Phage have the ability to lyse (kill) their bacterial host and this genetic mechanism has led to the development of phage therapy, a method used to treat infections. **This bioreactor would allow students to model the effect of phage therapy in addition to other treatments, such as antibiotics.** These advanced experiments allow for creative experimental design and a true ownership of a research project. Both of these characters are important for a transformative education. Reports indicate that CURE-like programs result in a higher retention in the sciences. Overall, having the MBRA to process student samples for this course would provide a tool for us to use to scale-up, serve more students, and make more discoveries.

2. **The second introductory lab course that will benefit from the MBRA is focused on the global challenge of soil.** In terms of food security for our planet, soil is the limiting factor. It is a grand challenge to determine how to recognize, produce, and maintain healthy soil. Soil is also a reservoir for carbon and therefore plays a fundamental role in determining the outcomes of another grand challenge, climate change.

   One of the areas in soil research that is not well-understood are the soil microbial ecosystems. We do not know which organisms are present in our soils or the roles that they play in nutrient recycling. Using this challenge, we are training introductory biology students to think like scientists and develop research skills that will address the soil microbiome. The recycling of nutrients and the health of the soil determines the health of plants and ultimately all animals which rely on plants for the basis of the food chain. Baylor students focus on the biology of a large and diverse group of single celled organisms, the ciliates. Ciliates are known to play an essential role in the soil ecosystem as predators of bacteria and other microorganisms. Due to the difficulty of culturing and classifying the ciliates, only around 3500 species have been described, but their specific role in the ecosystem is unclear. As next generation metagenomic techniques have improved, interest has grown in determining “who” is in the soil and ultimately what role are they playing in the ecosystem. The mini bioreactor will allow students to explore multiple samples and work on the most challenging questions by analyzing the environmental DNA on a large scale. The opportunity to provide this level of research to our students is exciting and will open doors for more transformational learning experiences.

3. Over 100 undergraduates participate in research under a mentor in the Department of Biology. **Almost every area of research has a bioinformatic component and this MBRA would allow undergraduates to participate in this community level of analysis.** It is a perfect system to consider for examining the

biological component of water quality, sewage effluent, and the dynamic microbiota of any system.

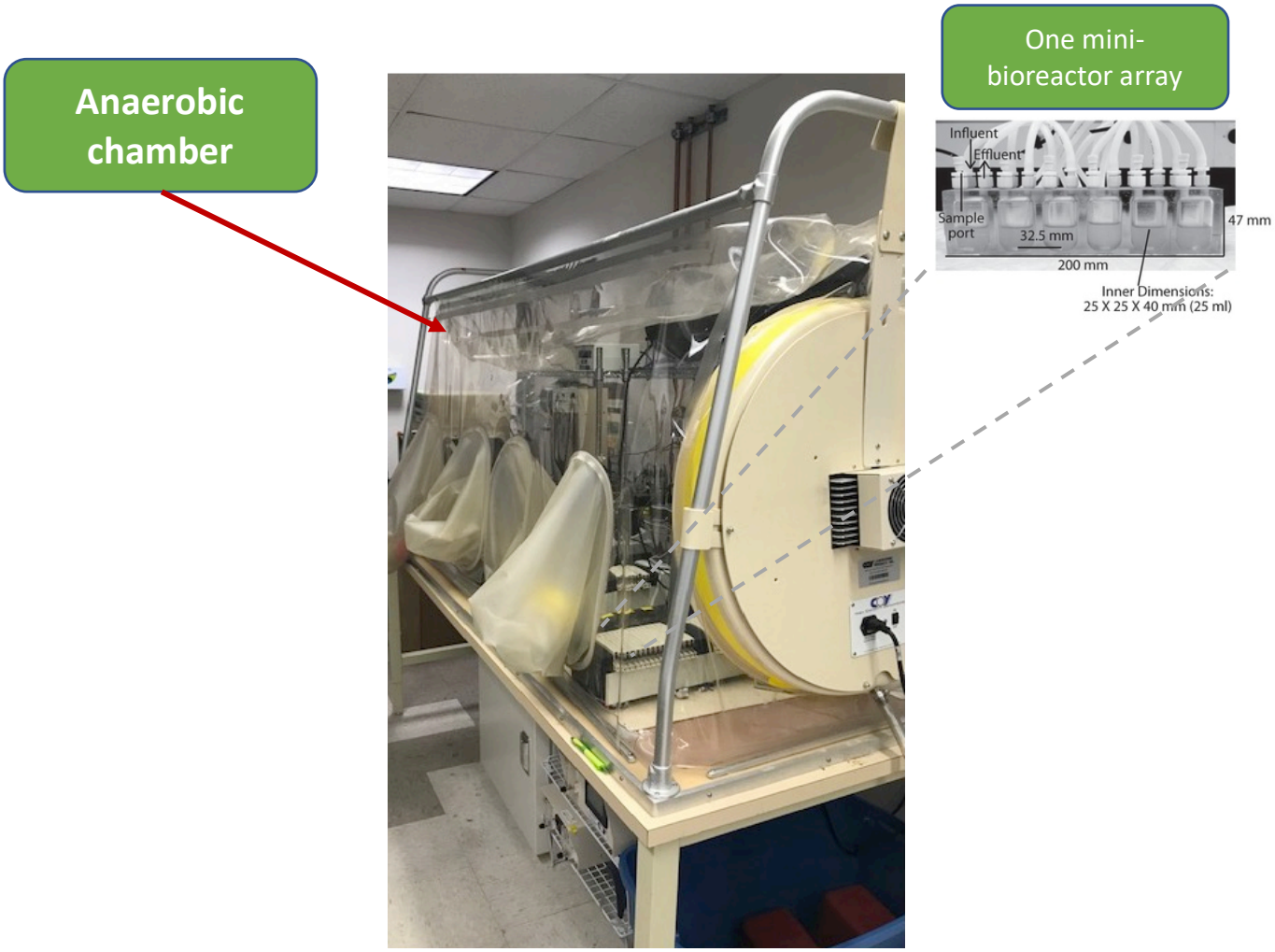
FACILITIES AND EQUIPMENT

This proposal seeks to purchase and establish a new piece of equipment at Baylor University, the mini-bioreactor arrays – MBRAs.

If this proposal is funded, it has been agreed (Drs. Meohnke and Chambliss) that space within the Molecular Bioscience Center in the Baylor Sciences Building would be utilized making this available to a larger user base. This would also allow Baylor to invest in this system as a core facility at a later date.

The MBRA system would be housed within a 4 ft x 2.5 ft **Whitley Workstation A35** anaerobic chamber, and would be fed through peristaltic pumps and three gas lines to allow bacteria to grow and propagate. Each system can house 28 arrays.

A similar larger MBRA system currently is in operation at Baylor College of Medicine in the laboratory of Robert A. Britton, PhD. **Pictured below:**



BAYLOR UNIVERSITY
Office of Sponsored Programs and Contracts
Detailed Budget

Proposal No.:	Illuminate		Sponsor: Baylor		Date: 3/27/19
Project Title:	Decoding the Gut Microbiome Response to Dietary Fiber Types Using Model Gut Communities				
Project Director(s):	Dr. Leigh Greathouse				Family & Consumer Scie
Project Dates:	8/1/19	to	7/31/22		
		Year 1	Year 2	Year 3	Total
		8/1/19-7/31/20	8/1/20-7/31/21	8/1/21-7/31/22	
Personnel:					
Postdoctoral fellow (0.5 FTE)		0	28,313	28,313	56,626
Graduate Student (GRA)		24,000	24,000	24,000	72,000
Total Personnel (S/W)		24,000	52,313	52,313	128,626
TOTAL SALARIES AND FRINGE:		24,000	52,313	52,313	128,626
Subcontract:					0
Nicholas Chia (Mayo)			13,853	13,853	27,706
Consultant/Services:					
Microbiome sequencing (BCM)		0	25,000	25,000	50,000
Metabolomics/SCFAs (Metabolome)		0	10,000	10,000	20,000
Total Consultant/Services		0	35,000	35,000	70,000
Training:		5,048	0	0	5,048
Training on MBRA system and setup					
Supplies:					
MBRA Equipment		89,208	0	0	89,208
qPCR reagents (DNA extractions kits, PCR master mix, plates)		0	6,250	6,250	12,500
Labware		14,203	0	0	14,203
Consumables (gas, gloves, tips, media, labware)		14,395			14,395
Total Supplies:		117,806	0	0	130,306
Travel		0	0	4,500	4,500
Total Direct Cost (TDC):		**146,854**	**87,313**	**91,813**	**366,186**
Total modified direct cost base		146,854	92,513	91,813	**0**
Indirect Cost:	**38.5% MTDC**	0	0	0	0
Total Cost:		**146,855**	**87,313**	**91,814**	**366,186**

DETAILED EQUIPMENT BUDGET

	THREE YEAR TOTAL	CURRENT FISCAL YEAR BUDGET YEAR 2019-2020 YEAR 1	2020-2021 YEAR 2	2021-2022 YEAR 3
DEPRECIATION				
Anaerobic chamber 1 (8 yr useful life)	$11,250	$3,750	$3,750	$3,750
Anaerobic chamber 2 (8 yr useful life)	$3,750			$3,750
Peristaltic pump - sets A-D (8 yr useful life)	$13,125	$4,375	$4,375	$4,375
Peristaltic pump - sets E-H (8 yr useful life)	$4,375			$4,375
Biosafety cabinet (8 yr useful life)	$3,187	$1,063	$1,062	$1,062
Ultra low temp freezer 1 (8 yr useful life)	$6,000	$2,000	$2,000	$2,000
Portable autoclave (8 yr useful life)	$3,705	$1,250	$1,250	$1,205
PCR machines (8 yr useful life)	$3,500		$1,750	$1,750
Microplate spectrophotometer (8 yr useful life)	$1,276		$638	$638
CO2 incubator (8 yr useful life)	$1,300		$650	$650
Total Depreciation:	$40,218	$8,688	$11,725	$19,805
SMALL EQUIPMENT				
Refrigerator/freezer	$515	$515.00		
Tube Vortexer	$250	$250.00		
60-spot stir plates (2)	$6,800	$3,400.00		$3,400.00
Single channel stir plates (1 unheated, 1 heated)	$561	$560.55		
Plate vortexer w/ adapter	$3,000		$3,000.00	
Balance	$260	$260.00		
Analytical balance	$350	$350.00		
Microwave	$170	$170.00		
Printer	$130	$130.00		
Electorphoresis equipment	$3,000		$3,000	
Total Small Equipment:	$15,036	$5,636	$6,000	$3,400
LABWARE				
Bioreactor strips (8)	$7,188	$2,396	$2,396	$2,396
Fittings, bottles, caps, beakers, flasks	$7,015	$3,515	$500	$3,000
2L bottles (2 cases of 10)		$962		
1 L bottles (1 cases of 10)		$176		
500 ml bottles (1 case of 10)		$143		
125 ml bottles (2 cases of 10)		$216		
2L erlenmeyer flasks (4)		$185		
1L glass beaker (Case of 6)		$166		
250 ml glass beaker (Case of 12)		$117		
stir bars (3 8 mm X 0.62 in;3 12.7 mm X 3 in)		$98		
mini stir bars (40 3X8 mm)		$104		
eppendorf racks (5)		$66		
2 hole bottle caps		$1,192		
10 1/4-28 mm thread to barbed male adapte		$90		
Total	$14,203	$9,336	$2,896	$5,396
CONSUMABLES				
Specialty Gases (10 Anaerobic, 10 Nitrogen, 2 CO2, 2 Blood gas)		$882		
Media	$354	$354		
Brain Heart Infusion		$224		
ATCC Vitamin Mix		$65		
ATCC Trace Mineral Mix		$65		
Plasticware (Pipet tips, disposal pipets, inoculatin	$3,973	$3,973		
Deep well plates		$394		
U bottom 96-well plates		$259		
96-well plate lids		$183		
PCR film		$79		
Tips		$2,778		
1.7 ml tubes		$280		
petri plates		$39		
50 ml conical tubes		$145		
serological pipets		$69		
gloves		$114		
biohazard bags		$395		
Sharpies		$24		
Aluminum foli		$10		
Lab tape		$25		
Paper		$15		
Electronic lab notebook		$240		
Freezer boxes		$400		
Peristaltic tubing & fittings	$3,732	$3,732		
Replacement Catalysts (2 yr/chamber)	$2,448	$612	$612	$1,224
Replacement Oxygen sensors (1 yr/chamber)	$2,732	$683	$683	$1,366
Replacment Hydrogen sensors (0.5 yr/chamber)	$1,156	$289	$289	$578
Total Consumables	$14,395	$16,328		
TRAINING				
Travel to learn new cultivation techniques			$4,000.00	
DEPARTMENTAL SERVICE CHARGES				
Total S & E for Departmental Service Charge (Year 1)		$39,987		
Total S & E for Departmental Service Charge (Year 2)		$20,621		
Total S & E for Departmental Service Charge (Year 3)		$28,601		
Total System Startup		**$89,208.53**		

Auchtung et al. Microbiome (2015) 3:42
DOI 10.1186/s40168-015-0106-5

Microbiome

RESEARCH Open Access

Cultivation of stable, reproducible microbial communities from different fecal donors using minibioreactor arrays (MBRAs)

Jennifer M. Auchtung[1], Catherine D. Robinson[2,3] and Robert A. Britton[1*]

Abstract

Background: Continuous-flow culture models are one tool for studying complex interactions between members of human fecal microbiotas because they allow studies to be completed during an extended period of time under conditions where pH, nutrient availability, and washout of waste products and dead cells can be controlled. Because many of the existing well-validated continuous-flow models are large and complex, we were interested in developing a simpler continuous-flow system that would allow microbial community dynamics to be examined in higher throughput while still maintaining complex microbial communities. To this end, we developed minibioreactor arrays (MBRAs), small volume bioreactors (15 ml) that allow simultaneous cultivation of up to 48 microbial communities in a single anaerobic chamber.

Results: We used MBRA to characterize the microbial community dynamics of replicate reactors inoculated from three different human fecal donors and reactors seeded with feces pooled from these three donors. We found that MBRA could be used to efficiently cultivate complex microbial communities that were a subset of the initial fecal inoculum (15–25 % of fecal OTUs initially observed). After an initial acclimation period of approximately 1 week, communities in each reactor stabilized and exhibited day-to-day variation similar to that observed in stable mouse fecal communities. Replicate reactors were predominately populated by shared core microbial communities; variation between replicate reactors was primarily driven by shifts in abundance of shared operational taxonomic units (OTUs). Consistent with differences between fecal donors, MBRA communities present in reactors seeded with different fecal samples had distinct composition and structure.

Conclusions: From these analyses, we conclude that MBRAs can be used to cultivate communities that recapitulate key features of human fecal communities and are a useful tool to facilitate higher-throughput studies of the dynamics of these communities.

Keywords: Human microbiome, Cultivation, Bioreactors, MBRA, Microbial communities

Background

The gastrointestinal microbiome plays an important role in health and disease (reviewed in [1–8]). Although many insights about the role of the microbiota have been gained by studying microbial community association with the human host (e.g., [9–11]), the availability of less complex models of the microbiota (i.e., conventional and humanized animal ([12–16] and in vitro models [17–22]) have also played an important role in elucidating the roles of the microbiota.

Continuous-flow culture models are beneficial for studying the complex interactions between members of the host microbiota in vitro because they allow for studies to be completed during an extended period of time under conditions where pH, nutrient availability, and washout of waste products and dead cells can be better controlled (reviewed in [23–25]). Although there are several well-studied and validated in vitro models of human microbiota (e.g., Simulator of Human Intestinal Microbial Ecosystem (SHIME, [21, 26, 27]), the TNO gastrointestinal model (TIM-2, [19, 28]) and the three-stage compound

* Correspondence: Robert.Britton@bcm.edu
[1]Center for Metagenomics and Microbiome Research, Department of Molecular Virology and Microbiology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
Full list of author information is available at the end of the article

continuous culture system [18, 20]), we were interested in developing a simpler, higher-throughput continuous culture system for human fecal communities.

To this end, we developed minibioreactor arrays (MBRAs, [29]). MBRAs were strips of six replicate bioreactors (each with a 15-ml operating volume) that were positioned on a 60-position magnetic stir plate. Continuous-flow was controlled by two 24-channel peristaltic pumps. Because of their relatively small size and simplistic design, up to 48 reactors could be run simultaneously in a single anaerobic chamber, thereby reducing the time and cost needed to evaluate multiple experimental perturbations to microbial communities.

Previously, we demonstrated that fecal microbial communities established in our MBRAs recapitulate one important aspect of healthy microbial communities—the ability to resist colonization by *Clostridium difficile* unless perturbed by antibiotics [29]. Further, these MBRA communities revealed differences in physiology between *C. difficile* epidemic strains that were supported by experiments in a humanized microbiota mouse model of *C. difficile*, providing additional support for the applicability of this model [29].

These initial studies primarily focused on *C. difficile* physiology in the context of disrupted MBRA communities and did not more broadly characterize the types of communities that could be cultivated. A more thorough characterization of the types of communities that could be cultivated in unperturbed MBRA was needed to evaluate the suitability of this platform for other studies characterizing microbial community dynamics and function. Therefore, we followed changes in microbial community structure over time in replicate MBRAs inoculated from three different fecal donors as well as MBRAs inoculated from a pool composed of these three donors.

We found distinct communities could be cultivated from each donor. Community composition stabilized within ~7 days of flow (~21 reactor turnovers based upon the 8-h retention time of MBRAs) to contain a core set of 40–45 operational taxonomic units (OTUs; clustered with ≥97 % ANI across the V4 region of the 16S rRNA gene) across replicate reactors from the same donor. These core set of OTUs contained ~65–95 % of the total sequences. Cultivation resulted in restructuring of the starting fecal communities, with modest decreases in Firmicutes, Actinobacteria, and unclassified bacteria coupled to increases in Bacteroides, Proteobacteria, and Verrucomicrobia. From our studies, we conclude that communities cultivated in MBRAs recapitulate key features of human fecal microbiota and that MBRAs are a useful tool to facilitate higher-throughput studies of the dynamics of these communities.

Results and discussion

Diverse microbial communities can be cultivated in MBRA

We inoculated triplicate reactors with fecal samples from one of three healthy donors (donor A, donor B, or donor C) or six replicate reactors with an equal mass of fecal sample pooled from each of the three donors (pool). After an initial acclimation period (16 h), we collected samples (day 1) and initiated continuous flow operation of the reactors with an 8-h retention time. We then collected samples from all communities daily for 20 additional days (days 2–21). We monitored changes in microbial communities by amplifying and sequencing the V4 region of the 16S rRNA gene from these samples as well as from the initial fecal inocula (day 0). We then quality-filtered the data and clustered the sequences into OTUs with ≥97 % average nucleotide identities before further analyses.

We found that cultivation in the MBRAs supported growth of diverse microbial communities. Microbial diversity, as measured by either the Inverse Simpson Index (Fig. 1a) or Shannon Index (see Additional file 1A for graph), was similar between the starting fecal inocula and the MBRA communities on days 2–21 in culture. Cultivation resulted in a decrease of the overall number of OTUs by approximately twofold relative to the starting fecal inocula (Fig. 1b); this was primarily due to loss of low abundance OTUs (OTUs with one sequence in the fecal samples, see Additional file 1B for graph). Consistent with a decrease in the number of low abundance OTUs, the overall evenness of OTU distribution increased approximately two- to threefold in the MBRA communities relative to the starting fecal inocula (Fig. 1c).

We observed that both microbial diversity and evenness spiked on day 1, after 16 h of incubation in medium and prior to the initiation of flow. This transient burst in diversity and evenness likely reflects the unique nature of the sample, which contained those community members that had begun growing as well as those community members that will not grow, either because they were non-viable or could not be cultivated under these conditions. (The method that we used to measure community composition, amplification of the 16S rRNA gene, does not distinguish live from dead cells.) However, once flow was initiated, non-viable and non-growing strains were lost by dilution and turnover within the MBRA, and the remaining community members were those capable of growth in the distinct MBRA communities.

MBRA cultivation impacts community composition and structure

We next examined the impact of cultivation on the composition and structure of MBRA microbial communities. We determined the relationships between communities

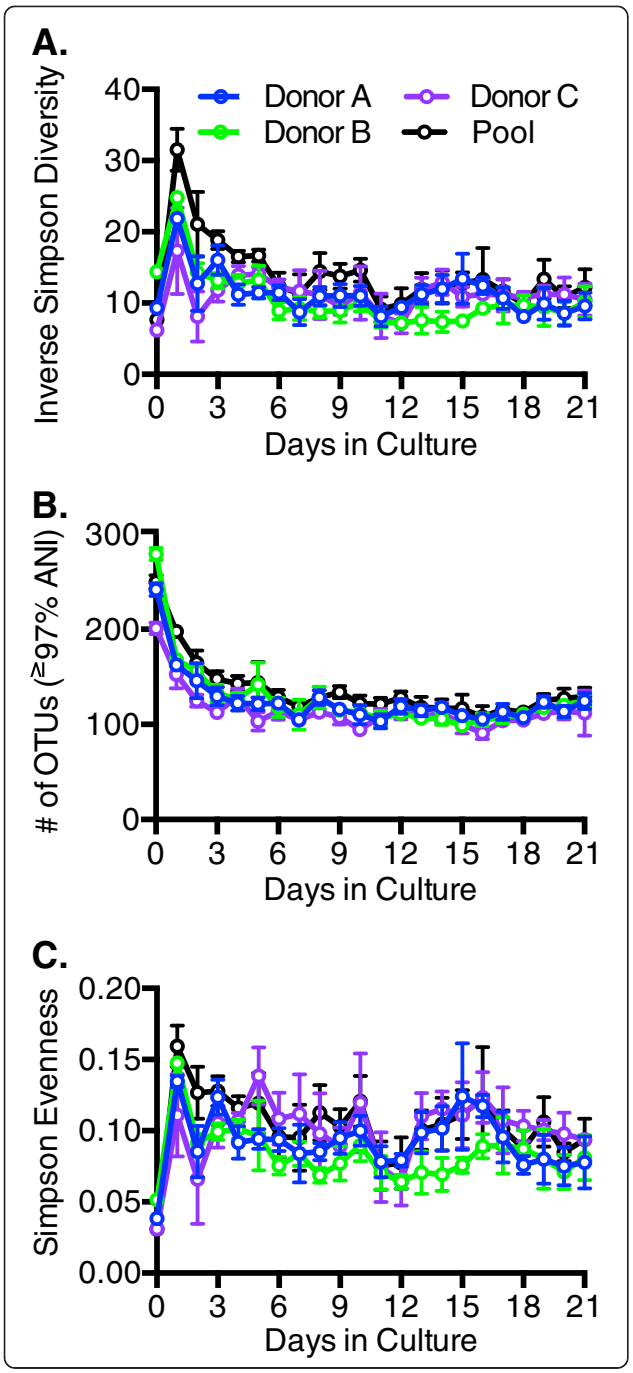


Fig. 1 Impact of MBRA cultivation on microbial diversity. Microbial diversity of triplicate MBRA communities inoculated with one of three donor fecal samples (Donor A, *blue circles*; Donor B, *green circles*; Donor C, *purple circles*) or in six replicate MBRA communities inoculated with an equal mass of all three donor fecal samples (Pool, *black circles*) was determined by sequencing the V4 region of the 16S rRNA gene from samples collected daily over 21 days in culture. Microbial diversity (Inverse Simpson, **a**), total number of OTUs (**b**), and evenness of OTU distribution (Simpson Evenness, **c**) was calculated from OTUs (≥97 % average nucleotide identity (ANI)) that were randomly subsampled to 10,000 sequences over 100 iterations. The mean value for the replicate reactors as a function of time in culture is plotted (day 0 = fecal inoculum; *error bars* represent standard deviation of the mean)

present in each sample using two similarity measures, Bray-Curtis similarity (Fig. 2a; Bray-Curtis similarity = 1-Bray-Curtis Dissimilarity), which calculates community similarity as the ratio of sequences in shared OTUs to total sequences, and Sorenson similarity (Fig. 2b; Sorenson similarity = 1-Sorenson Dissimilarity), which calculates community similarity as the ratio of shared to total OTUs [30]. We visualized the relationships between samples with non-metric multi-dimensional scaling (NMDS). As anticipated from previous studies (e.g., [31]), each individual fecal community was distinct. Although cultivation resulted in shifts in microbial composition compared to the starting fecal inocula (discussed more below), MBRA communities from each fecal donor (A, B, or C) rapidly formed distinct communities (Fig. 2a, b), with significant differences between communities inoculated from different fecal donors present by day 2 in culture.

The significance of differences detected between communities on days 2–21 were evaluated with two non-parammetric tests of community similarity, analysis of similarities (ANOSIM) [32] and PERMANOVA [33]. ANOSIM compares the rank similarities of samples within and between groups, generating an R-statistic varying from 1 (rank similarities within all replicates of a group are more similar than rank similarities between groups) to 0 (replicates within and between groups have similar rank similarities). Statistical significance is determined by comparing the R-statistic to the null distribution of R, which is calculated from iterations of randomly permuted data. PERMANOVA evaluates difference in community composition by calculating the differences within and between groups based upon the sum of squared distances from the centroid. The magnitude of the differences within and between groups is reflected in the pseudo-*F* value (*F* value >>1, reflects increasing differences in community composition), and the significance of this value is determined by random permutation. Using both tests for differences in community distributions, we observed that communities were separated with high significance ($p = 0.001$ based upon 999 permutations) and that this separation was

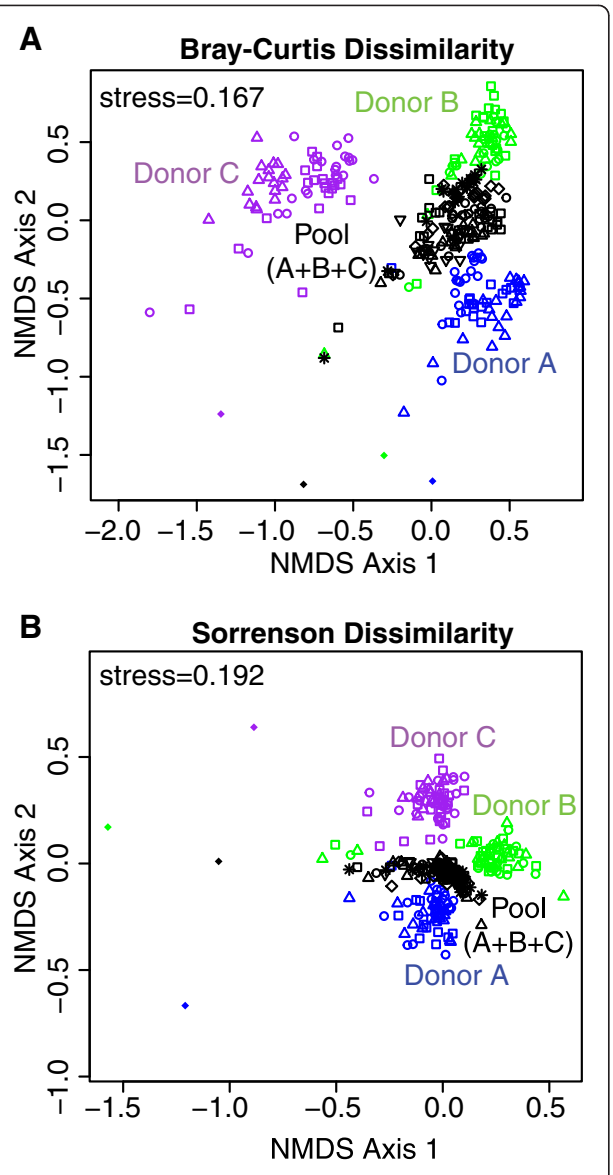


Fig. 2 Impact of MBRA cultivation on community composition and structure. Using the data described in Figure 1, we determined pairwise relationships between samples from MBRA communities inoculated with different fecal samples and their respective fecal inocula using (**a**) Bray-Curtis and (**b**) Sorenson dissimilarity measures and plotted this data with non-metric multi-dimensional scaling. Fecal samples = solid diamonds; MBRA communities = *open symbols*, with replicate 1 = *squares*, replicate 2 = *circles*, replicate 3 = *triangles*, replicate 4 = *diamond*, replicate 5 = *inverted triangle*, replicate 6 = *asterisks*; donors A, B, C, and the pool = *blue*, *green*, *purple*, and *black*, respectively. The stress for each NMDS plot is indicated

typically large (ANOSIM *R* values >0.89 and PERMANOVA F-statistics >55, Table 1).

MBRA communities inoculated with the pooled samples formed distinct communities. Comparison of differences between the communities formed from the pooled fecal sample to communities formed from donor A, B, or C also demonstrated that the donor C communities were less similar to the pooled communities than either donor A or donor B communities. These data indicate that the a larger proportion of community members present in donor C may not be able to compete with community members present in donor A or B during cultivation in MBRA.

Table 1 Evaluation of variation between MBRA communities inoculated with different fecal samples by analysis of similarities (ANOSIM, [32]) and permutational multivariate analysis of variance (PERMANOVA, [33])

	Bray-Curtis similarity			Sorenson similarity		
	ANOSIM R-statistic	PERMANOVA		ANOSIM R-statistic	PERMANOVA	
		F value	R^2		*F* value	R^2
Donor A-donor B	0.94	89.5	0.44	0.9	54.8	0.32
Donor A-donor C	0.99	106.7	0.49	0.98	76.4	0.4
Donor B-donor C	0.99	115.0	0.5	0.93	61.6	0.35
Donor A-pool	0.58	40.9	0.19	0.65	35.1	0.17
Donor B-pool	0.57	48.8	0.21	0.74	39.2	0.18
Donor C-pool	0.97	106.9	0.38	0.9	68.8	0.28

All *p* values were 0.001 based upon 999 permutations

MBRA cultivation leads to stable microbial communities

Community stability can be challenging to define, as one must decide how much variation is acceptable in a stable community. One previously reported measure for stability [34] that we applied to our data was to plot the mean Bray-Curtis similarity for each day's sample relative to all other days in culture as a function of time in culture (Fig. 3a). From this plot, we observed that mean similarity values increased rapidly over the first 3 days in culture (slope of line from day 0 to day 3, 0.13 ± 0.04 (mean ± SD)), continued to increase at a reduced rate through day 7 (Slope of line from day 4 to day 7, 0.019 ± 0.01), and plateaued around day 8 in culture (slope of line from day 8 to day 11, −0.009 ± 0.01). More precisely, the inflection point of each curve (i.e., point where slope transitioned from positive to negative) was identified by determining the slopes of each line with three-point sliding windows; the inflection point varied from day 8 to day 12, with a median of 8 and a mean of 8.6. By day 8 in culture, communities had experienced 7 days of continuous flow (~21 reactor turnovers with 8-h retention time). A second measure of stability [35], calculating the similarity between reactors as a function of increasing days in culture, demonstrated similar stabilization dynamics (Fig. 3b).

To better resolve the differences in community structure in stable bioreactor communities, we re-calculated the mean Bray-Curtis dissimilarity for each day's sample for days 8–21 and plotted this as a function of time (see Additional file 2). We observed that the mean Bray-Curtis similarities from day-to-day ranged from 0.58 to 0.86, with a mean of 0.74 ± 0.05.

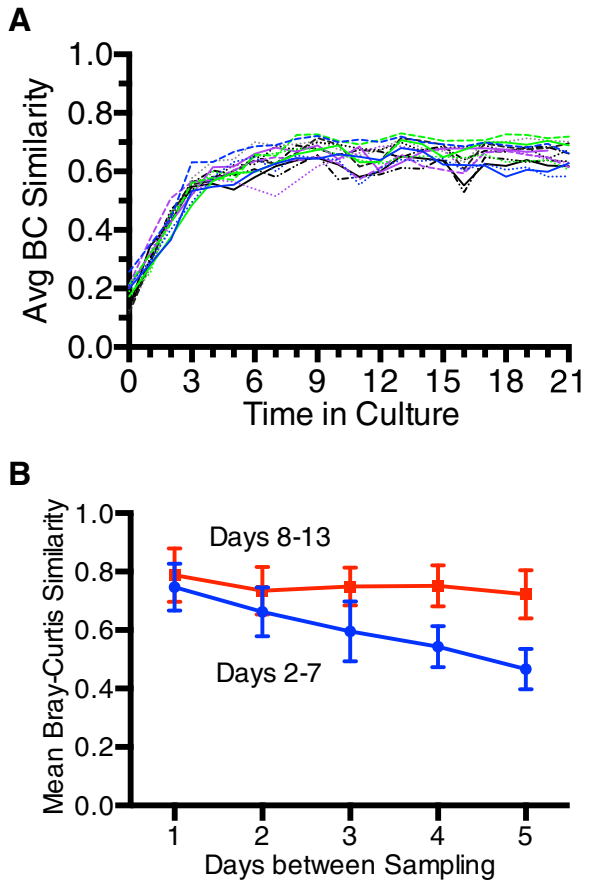


Fig. 3 Stabilization of MBRA microbial communities (**a**). MBRA community stability was assessed by plotting the average Bray-Curtis (BC) similarity between each daily bioreactor sample and other days in culture as a function of time in culture. The point at which reactors reached stability was defined as the inflection point of the curve and varied from day 8–12 (median = day 8). **b** The mean Bray-Curtis similarity (± standard deviation) for samples at increasing time intervals were calculated for all reactors over the indicated time intervals (days 2–7 (transitioning communities) and days 8–13 (stable communities)) and plotted as a function of days between samples. Statistical testing of each time interval with an unpaired student's *t* test demonstrated that the differences in similarity observed between transitioning (days 2–7) and stable (days 8–13) were significant ($p \leq 0.03$)

Previous studies have shown that variations in sampling, extraction, and sequencing can lead to introduction of artificial variation among samples (e.g., [35, 36]). In order to determine the amount of technical variation present in our data, we performed duplicate sequencing of three identical samples (described in the "Methods" section) and analyzed as described above. We found that the mean Bray-Curtis similarity between these technical replicates was 0.87 ± 0.04 (Additional file 2, red line). Thus, the mean variation in stable reactor communities was twice that observed in technical replicates; this twofold difference correlates to increased variation of sequence distribution of ~13 % in stable reactor communities relative to the controls.

Variation in microbial communities increases between replicate reactors

Having identified the window in which communities stabilize within each reactor (days 8–21), we next wanted to examine how these communities varied across replicates from the same fecal donor. We determined the mean similarities between replicate reactors inoculated from the same donor on days 8–21 and compared this to the mean similarities observed within each reactor from day-to-day on days 8–21 (Table 2). We observed that Bray-Curtis similarities decreased ~1.7–1.8-fold between replicate reactors. (Mean similarity values ranged from 0.54 ± 0.07 to 0.61 ± 0.08.) This divergence occurred primarily within the first 5 days of cultivation and did not increase significantly once individual reactors had stabilized (see Additional file 3 for graph of similarities between replicate reactors over time). In contrast, Sorenson similarity decreased ~1.1–1.2-fold between replicate reactors (Table 1; mean similarity values from 0.67 ± 0.04 to 0.70 ± 0.04). Although these decreases in Bray-Curtis similarities were modest, these data indicate that changes in the relative abundance of shared OTUs is one potential mechanism that led to differences between replicate reactors.

Comparison of variation in stable reactor communities to stable mouse communities

To gauge how MBRA community variation compares with another experimental model of gastrointestinal community dynamics, we compared our data to that published by Schloss and colleagues [37], who examined fecal microbial community stabilization in mouse microbial communities post-weaning. In this work, they observed that mouse communities exhibited the greatest variation in community structure in the first 9 days post-weaning and that community structure had stabilized by days 141–150 post-weaning. Therefore, we compared community dynamics in our stable MBRA communities (days 8–21) to differences in community dynamics within each mouse from day-to-day and between different mice over time. Although these models of GI community dynamics have several differences that could impact community dynamics (murine vs human microbiota; presence of host to provide different nutrients, niches, and selection from immune system), we were interested in examining the similarities and differences between community dynamics in these disparate models. We found that community dynamics were quite similar between mice and MBRAs.

Because their original work analyzed stability with a different dissimilarity measure (θ_{YC}) and used data

Table 2 Comparison of mean Bray-Curtis and Sorenson similarities for OTUs present in MBRA communities on days 8–21

Reactor type	Within reactor[a]	Between replicates[b]	Between reactor types[c]			
			A	B	C	Pool
Bray-Curtis						
Donor A	0.73 ± 0.08	0.54 ± 0.07	NA	0.29 ± 0.06	0.26 ± 0.05	0.47 ± 0.08
Donor B	0.79 ± 0.07	0.61 ± 0.08	0.29 ± 0.06	NA	0.25 ± 0.06	0.46 ± 0.10
Donor C	0.73 ± 0.10	0.55 ± 0.07	0.26 ± 0.05	0.25 ± 0.06	NA	0.32 ± 0.07
Pool	0.72 ± 0.09	0.57 ± 0.07	0.47 ± 0.08	0.46 ± 0.10	0.32 ± 0.07	NA
Sorenson						
Donor A	0.74 ± 0.05	0.69 ± 0.04	NA	0.55 ± 0.05	0.52 ± 0.05	0.64 ± 0.04
Donor B	0.71 ± 0.05	0.67 ± 0.04	0.55 ± 0.05	NA	0.53 ± 0.05	0.61 ± 0.04
Donor C	0.73 ± 0.04	0.67 ± 0.04	0.52 ± 0.05	0.53 ± 0.05	NA	0.57 ± 0.04
Pool	0.74 ± 0.04	0.7 ± 0.04	0.64 ± 0.04	0.61 ± 0.04	0.57 ± 0.04	NA

[a]Mean ± SD across all replicates of similarities within each replicate reactor of the indicated fecal type
[b]Mean ± SD across all pairwise comparisons of similarities between replicate reactors of the same fecal type on days 8-21
[c]Mean ± SD across all replicates of similarities between reactors of the different fecal types on days 8-21

generated by pyrosequencing the V3–V5 region of the 16S rRNA gene, we reanalyzed a subset of their data, which was generated by Illumina sequencing of the V4 region and used to cross-validate a new dual-indexing sequencing approach for community analyses [37], using the methods described above. Further, as they observed neither litter, co-housing status, nor sex of the mouse significantly impacted community structure, we selected data from three male and three female mice and treated these mice as independent replicates for our analyses.

Day-to-day variation within each mouse

We calculated the mean Bray-Curtis similarity within mice with stable communities to be 0.79 ± 0.06 (see Additional file 4 for a table of mean dissimilarities in stable and unstable mouse communities). Sorenson similarity values were similar, with mean within mouse values of 0.76 ± 0.04. From these data, we conclude that stable individual MBRA communities exhibit similar day-to-day variations as those found in stable murine communities.

Variation between replicate mice

The variation in Bray-Curtis similarities between replicate mice with stable communities was 0.71 ± 0.05 (see Additional file 4), which is ~10–15 % lower than the similarity observed between replicate reactors (0.54 ± 0.07 to 0.61 ± 0.08 to, Table 2). In contrast, the Sorenson similarity values between replicate reactors (0.67 ± 0.04 to 0.70 ± 0.04, Table 2) and replicate mice (0.72 ± 0.04, see Additional file 4) were similar.

One potential contributing factor to the higher variation observed in replicate reactors is the higher abundance of Firmicutes in MBRA communities compared to the mouse communities (56 ± 13 % in MBRA; 28 ± 8 % in mice). Flores et al. [38] examined temporal stability in human fecal communities over time and found those subjects that exhibited higher variation in community structure over time also had a higher ratio of Firmicutes/Bacteroidetes than those subjects with lower variation in community structure. As is discussed in more detail below, we found that the distribution of OTUs belonging to the Firmicutes and Proteobacteria phyla were more variable in our samples than OTUs belonging to the Bacteroidetes.

Another possible contributing factor to the lower variation observed in mouse data is that communities colonizing these mice are highly adapted to co-existence in the murine GI tract, where selection can be imposed by interactions with host cells, other members of the microbiota, and nutrients from the host diet. In contrast, cultivation in MBRAs could allow organisms with functional redundancy under MBRA cultivation conditions to fluctuate stochastically during stabilization until stable communities are reached.

Composition of core MBRA communities and comparison to starting fecal inocula

To gain further insights into the composition and structure of MBRA communities, we identified those OTUs that were conserved across samples (i.e., core communities). We identified core communities on three different levels: OTUs found conserved from day-to-day over days 8–21 within each single reactor (individual core), OTUs shared from day-to-day over days 8–21 between replicate reactors of the same fecal donor (fecal type core), and OTUs common to all reactor communities over time (all MBRA core). We also compared the composition of these core communities to their starting fecal inocula.

Identification of core communities

OTUs were designated as members of a core community if they were present in at least 90 % of daily samples collected from a single reactor between days 8–21. Using these criteria, we found that the core communities maintained from day-to-day within each reactor varied from 55 to 72 OTUs (~50–60 % of total OTUs from each day sampled; Fig. 4a (individual cores)). These individual core OTUs contained 95–98 % of the sequences from each day sampled (Fig. 4b). When we determined the overlap between OTUs found in the core communities of replicate reactors of from the same fecal donor, we found that the core community shared between replicate reactors of the same fecal type was composed of 40–45 OTUs (~30–40 % of total OTUs, Fig. 4a (fecal type core)) and contained 66–95 % of sequences present in each reactor from each day sampled (Fig. 4b). Finally, we determined the overlap in core membership across MBRA replicates from all fecal types and found that this all MBRA core contained 12 OTUs comprising 18–48 % of total sequences from each reactor across every day sampled (Fig. 4; also see Additional file 5 for a table listing the 12 OTUs comprising the all MBRA core and their abundances across samples).

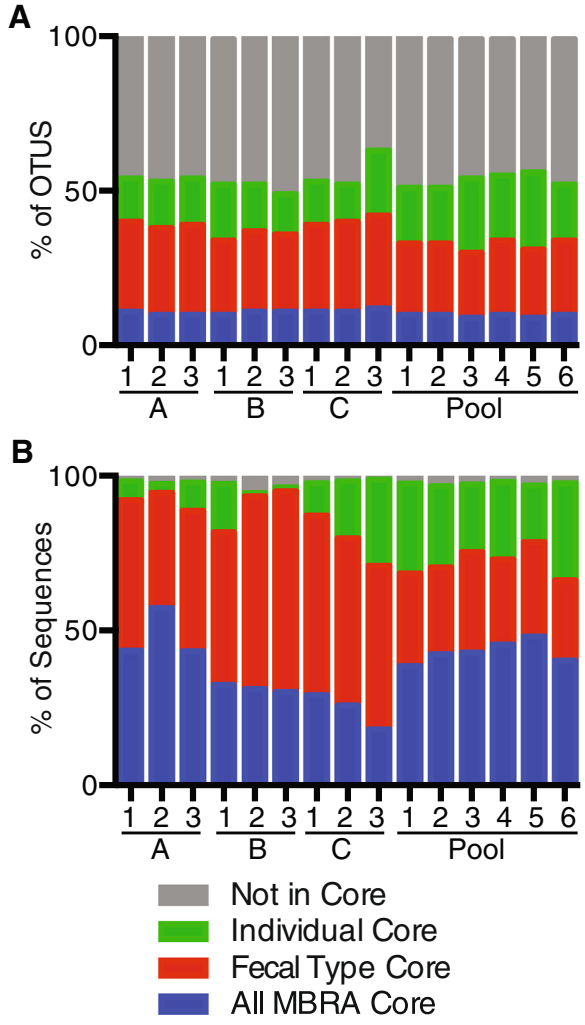


Fig. 4 Structure of MBRA core microbial communities. We designated those OTUs that were present in ≥90 % of daily samples over days 8–21 within each single reactor over time as members of the "individual core." Individual Core OTUs that were shared across replicates of the same fecal type were designated members of the "fecal type core" and those present in the core of all MBRA communities were designated "all MBRA core." We calculated the mean percent abundance of OTUs (**a**) and sequences (**b**) in each type of core from reactors on days 8–21 and plotted these values for each replicate (1–3 for donors A, B, or C; 1–6 for pooled fecal donor). Each core type includes those members also present in the broader core type (i.e., individual core = individual core + fecal type core + all MBRA core; fecal type core = fecal type core + all MBRA core)

Comparison of MBRA communities with fecal inocula

We also examined the phylogenetic distribution of OTUs present in the MBRA core communities and compared this with the starting fecal inocula. Because Individual Core OTUs contained 95–98 % of the MBRA sequences, we limited our analyses to these 144 OTUs. In addition, we included those 28 OTUs that were abundant in fecal samples (≥0.5 % of sequences) and were not present in any of the individual core communities (see Additional file 6 for a table listing the 28 OTUs abundant in fecal samples and absent from individual cores their abundances across fecal samples). Limiting our analyses to these 172 OTUs greatly simplified the amount of data while still representing the distribution of ≥94.5 % of the sequences from every sample.

At the phylum level, MBRA communities exhibited similar trends across replicate reactors and community types (Fig. 5a). Sequences classified as Firmicutes were the most abundant members of the community, ranging from 44–70 %, followed by Bacteroidetes (18–32 %), Proteobacteria (5–17 %), and Verrucomicrobia (3–8 %). At this level of analyses, donors A and B MBRA communities were also similar to their starting fecal inocula, whereas donor C and the pool MBRA communities were different. In the fecal inoculum for donor C, the ratio of Firmicutes/Bacteroidetes sequences was higher (80 % Firmicutes sequences/10 % Bacteroides sequences) than the ratio in the fecal inoculum from either donor A (61 % Firmicutes/27 % Bacteroidetes sequences) or donor B (fecal inoculum in the fecal inoculum from donor C (74 % Firmicutes/17 % Bacteroidetes). Growth under MBRA cultivation conditions led to a decrease in the ratio of Firmicutes/Bacteroides in reactor communities seeded from donor C (mean ratio in MBRA communities from donor C: 56 % Firmicutes/32 % Bacteroidetes), which was more similar to that observed for MBRA communities seeded with other fecal samples. The phylogenetic distribution of the pooled fecal inoculum was unexpected, with

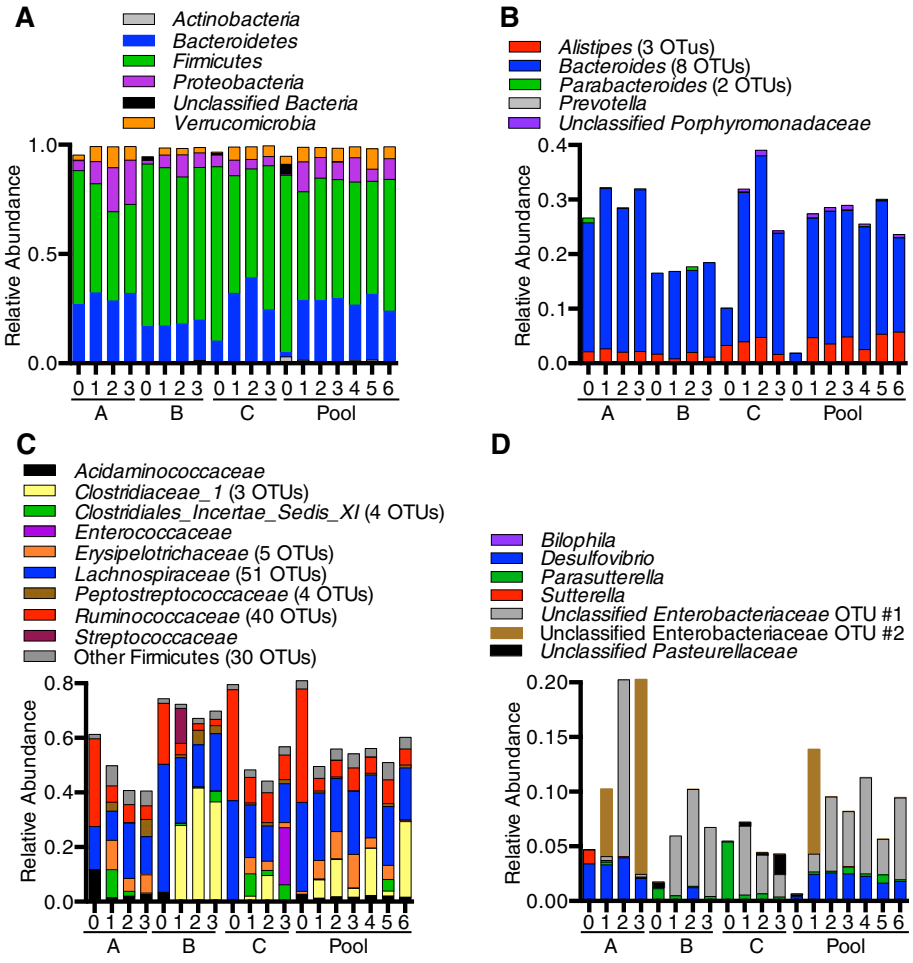


Fig. 5 Composition of MBRA core communities and comparison with fecal inocula. We analyzed the phylogenetic distribution of OTUs in the individual core communities and compared this with the phylogenetic distribution of the original fecal inocula. To provide better representation of the fecal inocula, those OTUs absent from the individual core communities that contributed at least 0.5 % of sequences to a fecal sample were also included in our analyses. As in Fig. 4, data present is the mean abundance for each OTU across days 8–21. Following determination of consensus classifications for each OTU (as described in Methods), we plotted the percent abundance of sequences in each phylum (**a**), genus within Bacteroidetes (**b**), family within Firmicutes (**c**), and genus within Proteobacteria (**d**). If a consensus classification for the phylogenetic level plotted could not be determined with confidence, the next highest classification assigned with ≥80 % is given preceded by the designation "unclassified." To simplify presentation of abundances in **c**, several families with low abundance across all samples were condensed in to the designation "other Firmicutes", which includes *Clostridiales Incertae Sedis XIII*, *Eubacteriaceae*, *Incertae Sedis XI*, unclassified Firmicutes, unclassified Clostridia, unclassified Clostridiales, and Veillonellaceae

only 2 % of sequences from Bacteroidetes, 81 % of sequences from Firmicutes, 3 % of sequences from Actinobacteria and 4 % of sequences from unclassified bacteria. The levels of Bacteroidetes were 5–14-fold lower than observed in any of the individual fecal samples used to generate the pool; whereas the levels of Actinobacteria and unclassified bacteria were 15–37-fold and 3–15-fold higher, respectively, than any of the individual fecal samples used to generate the pool. (Levels of *Firmicutes* (81 % of sequences), Proteobacteria (2 % of sequences), and Verrucomicrobia (4 % of sequences) were less than twofold different from at least one of the individual inocula.) This distribution could have been caused by heterogeneity in the individual fecal samples used for preparation of the pooled fecal sample and was not reflected on pooled MBRA communities, which looked more similar to other MBRA communities.

The variation between reactor communities and their fecal inocula was more apparent when OTU distributions within specific phyla were examined (Fig. 5b–d); these differences were primarily driven by members of the Firmicutes and Proteobacteria (Fig. 5c, d; also see Additional file 7 which presents Bray-Curtis similarities for MBRA communities and their fecal inocula based upon all OTUs as well as for OTUs partitioned by phyla). The largest diversity across reactors was observed

in the Firmicutes; this phylum was represented by 144 OTUs (38 genera from 16 families). Although Proteobacteria were less abundant members of the communities (4–20 % of sequences), they were represented by seven different OTUs across six genera. In contrast, OTUs within the Bacteroidetes phylum were primarily members of the Bacteroides genus (Fig. 5b); all Verrucomicrobia sequences were from a single *Akkermansia* OTU (classified using the Greengenes reference taxonomy as *Akkermansia muciniphila*). From these data, we conclude that stable MBRA communities represent a subset of their fecal inoculum. The percent of the initial fecal inoculum present in core communities vary (from as little as 25 % of sequences (15 % of OTUs) in donor B communities to as much as 49 % of sequences (22 % of OTUs) in donor A communities). The amount of overlap observed was dependent upon the composition of the community present in the initial inoculum and could reflect both differences in abundance of obligate anaerobes that were lost prior to cultivation as well as differences in ability to grow under our cultivation conditions.

Examination of core OTU dynamics over time in culture

Figure 5 presented the average abundance of taxa found in ≥90 % of the samples collected on days 8–21 but did not indicate how levels of these taxa varied over time. To evaluate variation over time, we used heat maps to visualize changes in abundance of core OTUs from the starting fecal samples throughout time in culture (days 1–21). Variation in core OTUs from donor A are shown in Fig. 6, whereas data from donors B and C are shown in Additional files 8 and 9.

Analysis of the data from MBRA communities seeded with donor A demonstrated three general trends. (1) Of OTUs present in the fecal sample, ~20 % were lost or decreased >5-fold during the first several days in culture (e.g., *Faecalibacterium*, *Roseburia*, some Bacteroides, and unclassified Lachnospiraceae; indicated OTUs are identified by red text in Fig. 6). (2) Of OTUs, ~10 % persisted in MBRAs from day-to-day at levels <3-fold different than those observed in fecal samples (e.g., *Flavonifracter*, *Akkermansia*, Bacteroides #4, Ruminococcaceae #5; indicated OTUs are identified by blue text in Fig. 6). (3) The remaining ~70 % of OTUs appear to increase upon cultivation in at least one replicate tested. Similar patterns of increase and loss of OTUs were seen across MBRA communities inoculated from donors B and C and the pooled fecal samples, although the organisms impacted varied across the different fecal donors. For example, *Roseburia* OTUs were lost from donor A and B communities upon cultivation but persisted in communities inoculated with samples from donor C (see Additional file 9, Roseburia OTU is in red type).

Included in the group of OTUs that increase upon cultivation were organisms classified as Enterobacteriaceae, which constituted 2–18 % of total sequences found across replicate reactors (mean abundance of two Enterobacteriaceae OTUs = 8 ± 5 % of sequences). Although many Enterobacteriaceae are facultative organisms, these organisms are unlikely to be respiring aerobically because oxygen levels in the anaerobic chamber are kept below 20 ppm. Organisms from the two Enterobacteriaceae OTUs could be respiring anaerobically, using alternative electron acceptors such as TMAO (known to produced in MBRAs; JMA and RAB, unpublished results) or fumarate (likely metabolic byproduct in reactor), or could be fermenting available carbohydrates.

Conclusions

In this paper, we demonstrated that MBRAs can be used to efficiently cultivate distinct communities from multiple fecal donors. Within the first week of cultivation, distinct microbial communities capable of metabolizing the available nutrients developed from the different starting inocula. Adaptation to growth in culture shifted the community structure. Although some community members persisted at similar abundances to the fecal inocula, other rare members of the inoculum increased in abundance and a subset of the initial inoculum was lost. This adaptation followed similar trends across replicates, with stable communities obtained by day 8 in culture (7 days with flow; ~21 turnovers).

Day-to-day variation within MBRA communities from single reactors were similar to those observed in stable mouse fecal communities. In contrast, replicate reactors from the same fecal donor exhibited slightly higher variation than was observed between replicate mice. This variation appeared to partly be driven by differing abundances of shared OTUs between replicates and could be indicative of functional redundancy of organisms under the MBRA cultivation conditions. This functional redundancy could allow abundances of OTUs to fluctuate stochastically until stable communities are reached. Further work will be needed to investigate functions of these different communities, although we know that communities formed from all three fecal donors as well as the pooled fecal sample are capable of resisting colonization by *C. difficile* in the absence of perturbation (JMA, CDR, and RAB, unpublished results).

The composition and structure of MBRA communities are similar to those reported in other in vitro models of human fecal communities [27, 28, 39], although the levels of Bacteroidetes are lower than are typically observed in many models. One potential factor that might lead to the lower levels of Bacteroidetes observed under our culture conditions is the low concentrations of fermentable carbohydrates present in our medium, which

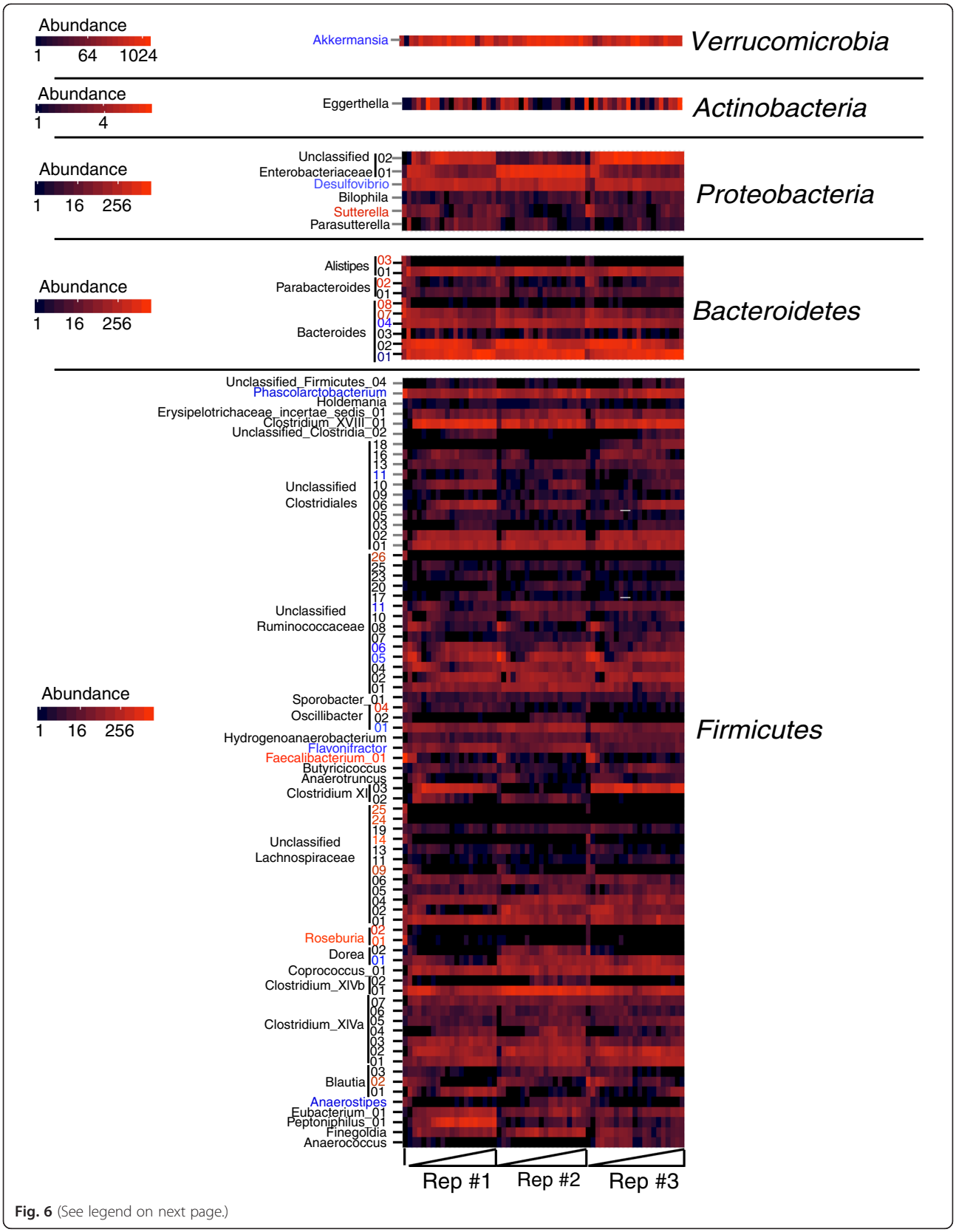


Fig. 6 (See legend on next page.)

(See figure on previous page.)
Fig. 6 Analysis of abundance of core OTUs identified in Donor A MBRA communities as a function of time in culture. We determined the abundance of the 94 OTUs that were identified as present in individual core communities for MBRAs inoculated with fecal donor A or were abundant in the donor A fecal sample as described in Fig. 5 and plotted the abundance of these OTUs in the fecal sample and over time in culture (days 1–21) across the three replicate reactors. Data are organized by phylum, with the lowest taxonomic classification assigned with confidence listed on the left hand side. Magnitude of shading is indicated on the figure and ranges from 1 to ≥256 sequences for Firmicutes, Bacteroidetes, and Proteobacteria. Abundance of Actinobacteria sequences range from 1 to 4 sequences; whereas the abundance of the single Verrucomicrobia OTU range from 1 to ≥1024 sequences. The line at the left end of the *x*-axis indicates the fecal sample. The *triangles* demarcate time in cultures for the different replicate reactors, with the first time point present on the left side for each replicate. Similar heat maps for donor B and donor C are available in Additional files 8 and 9

are a known substrate for Bacteroidetes (reviewed in [40]). However, as the levels of Bacteroidetes detected are similar to those present in the initial fecal inocula, we do not consider the existing levels of Bacteroidetes in MBRA communities to be of concern.

The choice of medium used to cultivate human fecal communities can have a significant impact on the communities that are formed (e.g., [40]). Although many of the existing fecal bioreactor media share some reagents, media can vary significantly in composition from one experimental setup to the next (see Additional file 10, a table comparing media composition from 11 different bioreactor studies). When evaluating media for use in our MBRA model, we compared different published media recipes, with specific emphasis on those models that supported dynamic interactions between human fecal communities and *C. difficile* (i.e., [20]). We also considered the results published by MacFarlane and colleagues [18], which demonstrated significant depletion of fermentable carbohydrates in the third (distal colon mimicking) vessel of their three-stage reactor model. Based upon these observations as well as our own preliminary studies with medium containing higher levels of carbohydrates, we chose to use a medium low in glucose (0.004 %, 220 μM), disaccharides (0.015 % (440 μM) each of maltose and cellobiose), and complex carbohydrates (inulin, 0.02 %; arabinogalactan, 0.01 %), in an attempt to more closely simulate conditions that might be encountered in the distal colon. (Although glucose and maltose are primarily absorbed in the small intestine, inulin, arabinogalacton, and cellobiose have been shown to be primarily fermented in the colon [41–43]). We also included Tween 80, a reagent commonly added to bioreactor medium (see Additional file 10, a table comparing media composition from 11 different bioreactor studies). Tween 80 can be used as a source of unsaturated fatty acids by several *Lactobacillus* species and has been shown to enhance growth [44] and protect from bile acid stress [45].

Although our model does not include a surrogate for mucosal surface as in M-SHIME, our phylogenetic distributions were more similar to the luminal contents of the M-SHIME model than to the original L-SHIME model [27, 39], with Firmicutes composing 44–70 % of the population. Further, we found that sequences of the known mucin-degrader *Akkermansia muciniphila* [46] were present in modest abundance in all stable reactor communities (~5 % of sequences). Further work will be needed to determine how MBRA cultivation conditions are supporting growth of *Akkermansia* in the absence of mucin.

One limitation to the communities cultivated in our MBRA is our inability to cultivate *Faecalibacterium prausnitzii* (11–36 % of sequences from fecal inocula) as well as several other less abundant *Clostridiales* species. Work from the Flint laboratory has demonstrated that both *F. prausnitzii* as well as certain members of *Clostridium* cluster XIVa require acetic acid as a cofactor for metabolism (reviewed in [40]), which is not present in our medium and is unlikely to be produced until a mature community is established. Further, *F. prausnitzii* does not grow well in medium with protein as the primary dietary substrate. Therefore, we may need to consider additional modifications to the medium to include a higher concentration of fermentable carbohydrates as well as a potential source of acetate during outgrowth to facilitate cultivation of *F. prausnitzii.*

A second limitation of our MBRA model is the inability to study interactions with host cells, both to assess how the microbiome impacts the host and to determine feedback of the host upon the microbiome. Although this limitation can be overcome by coupling MBRA studies with follow-up studies in humanized microbiota mice [29], it would also be beneficial to begin to interface MBRA communities directly with host cells in vitro. Platforms for interfacing microbial communities with tissue culture cells have been described [47, 48], including the HMI model for SHIME [47]. We anticipate that the relatively simple MBRA design could make it ideal for coupling with host cells, either individual cell lines or human intestinal enteroids or organoids [49, 50], thereby facilitating higher-throughput in vitro studies of microbiome/host interactions.

Methods

Fecal sample collection, preparation, and MBRA operation

Fecal samples from three healthy individuals were collected into sterile containers, sealed, and transferred to

an anaerobic chamber within 1 h of defecation. Samples were manually homogenized and subdivided into sterile vials, which were stored at −80 °C until use. Prior to MBRA inoculation, fecal samples were resuspended at 25 % *w*/*v* in anaerobic phosphate buffered saline in the anaerobic chamber, vortexed for 5 min, and centrifuged at 201×*g*. For the pooled sample, equal amounts of each fecal sample (by mass) were combined prior to vortexing.

In order to analyze the impact of freezing upon MBRA cultivation, one fecal donor (donor A) provided a second sample ~3 months post initial donation (donor A2). This sample was collected and transferred to the anaerobic chamber within 1 h of defecation. Following manual homogenization and subdivision into sterile vials, a portion of the sample was flash frozen in liquid nitrogen for 45 min (frozen), whereas the other sample was maintained in the anaerobic chamber until inoculation. Both fresh and frozen samples were then inoculated into triplicate reactors and analyzed as described below. Analysis of this data revealed that there was little impact on communities cultivated from frozen samples compared to freshly voided samples (see Additional file 11 for an NMDS ordination of the Bray-Curtis dissimilarities between these different samples as well tests of community similarity and dispersion).

MBRA were prepared for use as previously described [29] and inoculated with 4 ml of fecal slurry. Bioreactor medium was prepared as described [29], except that 1 g/L of taurocholic acid was replaced with 0.5 g/L of bovine bile, which was added prior to autoclaving. There were multiple reasons for substituting bovine bile for taurocholate. (1) Bovine bile is a complex mixture of bile salts as well as other constituents of bile (e.g., fatty acids, cholesterol, inorganic salts) and is more commonly used in medium for cultivation of human fecal communities than taurocholate alone. (2) Taurocholate was originally included in our medium to promote germination of *C. difficile* spores; subsequent studies have shown that bovine bile is sufficient to support germination under our reactor conditions (Auchtung and Britton, unpublished results). (3) Bovine bile is significantly less expensive than taurocholate (>10-fold lower cost).

After inoculation, fecal bacteria were allowed to equilibrate for 16–18 h prior to the initiation of flow. After equilibration, a 1-ml sample was removed (day 1 sample) and flow commenced at 1.875 ml/h (8-h retention time). Reactors were then sampled daily for 20 additional days (days 2–21). Cells were pelleted from samples by centrifugation at 21,000×*g*. Supernatants were discarded, and pellets were stored at −80 °C until further processed.

Ethics, consent, and permissions

Fecal sample collection was reviewed and approved by the Institutional Review Board from Michigan State University. All individuals donating samples provided informed consent prior to donation.

Sample preparation and sequencing

Previously, we had success amplifying *C. difficile* genes from samples that had been disrupted by bead beating without further purification [29]. Further, Flores et al. had also reported success with a direct amplification approach for higher-throughput analysis of 16S rRNA gene content from microbial samples [38]. We were interested in pursuing direct amplification in order to significantly reduce sample preparation time and costs. Therefore, we performed preliminary studies to compare sequences obtained from replicate samples prepared by direct amplification to those that obtained from samples from which DNA was extracted prior to amplification. These studies, which demonstrated robust reproducibility between duplicate samples prepared by direct amplification, are described in detail in Additional file 12: Supplementary Methods and Additional files 13, 14, and 15 (which present the data described in additional file 12).

We resuspended our samples in a 0.5-ml sterile water and transferred them to bead beating tubes. (Our bead beating tubes were prepared by transferring ~200 μl 0.1 mm silica beads (Biospec Products) and 100 μl sterile water into 2-ml screw cap tubes and autoclaving these tubes for at least 20 min prior to use.) Samples were homogenized in a mini-beadbeater-96 (Biospec Products) for 2 min, centrifuged at 8000×g for 1 min, then supernatants were transferred to new tubes, which were stored at -20 °C prior to amplification.

The V4 region of the 16S rRNA gene was amplified with primers F515/R806, using a dual-indexing approach (4 forward primer; 96 reverse primer). The 96-indexed R806 primers used were previously described ([51]; 806rbc0-806rbc96). The indexed F515 primers were essentially as described [37], except that we generated four barcodes that balanced the nucleotide composition at each position (ATCGATGG, TCACGACA, GGTATCTC, and CAGTCGAT) in place of those described by Kozich et al. Prior to PCR amplification, samples were diluted 1:100. The final 25-μl PCR reactions contained 4 μl of diluted template, 1× Phusion High Fidelity Buffer (New England Biolabs), 200 μM dNTPs (Promega), 10 nM primers and 0.225 units of Phusion DNA Polymerase (New England Biolabs). The amplification cycle consisted of an initial denaturation at 98 °C for 30 s, followed by 30 cycles of 10 s at 98 °C, 20 s at 51 °C, and 1 min at 72 °C. Successful amplification was verified by agarose gel electrophoresis of products. If samples failed to amplify, amplification with a new 1:100 dilution was attempted. If re-amplification failed, amplification was attempted with a 1:10 dilution of sample. Because we were unable to obtain amplification from the fecal slurries from donor A, B, C, or the pooled

sample at these dilutions, we extracted DNA as previously described [29] prior to amplification and used 4 μl of 10 ng/μl DNA in PCR reactions as described above. As discussed in more detail in Additional file 12, which compares extracted and amplified fecal samples to MBRA communities, comparing MBRA communities that were directly amplified to extracted fecal samples is one potential source of variation between our MBRA communities and fecal samples. However, sample preparation method prior to sequencing is likely not the primary source of variation between fecal and reactor communities as differences in sample preparation method in our control studies resulted in Bray-Curtis similarities of 0.64–0.78 (see Additional file 13 for the impact of sample preparation on Bray-Curtis and Sorenson similarity measures) as compared to the Bray-Curtis similarities observed between fecal and MBRA samples of 0.08–0.40 (mean = 0.17). Further, we were able to successfully amplify community DNA from the donor A2-fresh and frozen samples described in Additional file 11 without DNA extraction, yet observed similar levels of dissimilarity between these fecal inocula and their respective MBRA communities.

Three independent PCR replicates were pooled and cleaned up using AMPure beads as previously described [29]. Concentrations of purified DNA samples were determined with QuantIT (Life Technologies). Purified samples were pooled at equimolar ratios, and the quality of the pooled DNA was assessed by analysis on a Bioanalyzer High Sensitivity DNA Kit (Agilent). Prior to sequencing, DNA concentration was determined by amplicon-specific qPCR (Illumina Complete Kit, Kapa Biosystems). Samples were mixed with 3–7 % phiX DNA, and sequencing was performed at the Research Support Technology Facility (RTSF) with a MiSeq v2 Reagent kit on an Illumina MiSeq running MisSeq Control Software version 2.3.0.3. Sequencing was completed in two separate MiSeq runs.

Analysis of sequence data

Sequences were analyzed in Mothur versions 1.31, 1.33, and 1.34 essentially as described [37]. MiSEQ SOP version 28 March 2013 was used as a template (http://www.mothur.org/wiki/MiSeq_SOP). Forward and reverse reads were paired, quality-trimmed, aligned to a Silva 16S rRNA gene reference database, trimmed to ensure overlap to the same region of the 16S rRNA gene (position 534-786 of *Escherichia coli* 16S rRNA gene), and pre-clustered to clusters with ≥99 % identity as described. Potential chimeric sequences were identified with the mothur-implementation of uchime and removed. Sequences were then classified with the Bayesian classifier in mothur, using the mothur-formatted ribosomal database project version 9 database from August 2013. Sequences were clustered from a distance matrix using the average-neighbor algorithm in mothur. Taxonomic assignments for each OTU were determined in mothur and are the majority consensus taxonomic assignments for each sequence within the OTU.

The mean and SD of inverse Simpson diversity, Simpson evenness, Shannon diversity, and number of observed species were calculated in mothur from data randomly subsampled to 10,000 sequences over 100 iterations. A single iteration of subsampling to 10,000 sequences was used for determination of Bray-Curtis and Sorenson dissimilarity measures. Bray-Curtis and Sorenson dissimilarity values were calculated on untransformed data, both in mothur and with vegan package of R ([52]; Sorenson = Binary Bray-Curtis); when presented as similarity values, Bray-Curtis and Sorenson similarities = 1-dissimilarities. Non-metric multi-dimensional scaling was also performed in mothur and in vegan using the metaMDS function. NMDS Plots were generated in R from the metaMDS results. ANOSIM, PERMANOVA (ADONIS), and Betadispersion were calculated with the vegan package of R. Heatmaps were generated using the phyloseq implementation of NeatMap [53]. All other plots were generated in Graph Pad Prism v.6.

For the mouse stability studies, paired Illumina reads were downloaded from the Schloss lab website (http://www.mothur.org/MiSeqDevelopmentData.html). Sequences were processed through mothur as described above, then subsampled to 3500 sequences prior to calculation of beta-diversity measures. Data were from three female (F3, F4, and F7) and three male (M2, M5, and M6) mice on days 1–9 (unstable communities) and days 141–150 (stable communities).

Availability of supporting data

The sequence data described in this manuscript is can be accessed from the short read archive at NCBI (SRP059604).

Additional files

Additional file 1: Additional measures of MBRA diversity. Plots of Shannon Diversity and the number of OTUs containing at least one sequence.

Additional file 2: Stabilization of MBRA communities. Plot of Average Bray-Curtis (BC) dissimilarities on days 8-21 across replicate reactors.

Additional file 3: Similarity between replicate reactors increases over the first week of cultivation. Plot of mean BC similarities between replicate reactors over time in cultivation.

Additional file 4: Mean Bray-Curtis and Sorenson dissimilarities for OTUs in stable and unstable mouse communities. Table providing mean Bray-Curtis and Sorenson dissimilarities within individual mice and between replicate mice based upon shared OTU content.

Additional file 5: Relative abundance of all MBRA core OTUs across stable reactors (days 8–21; mean ± SD). Table providing the mean

relative abundance of OTUs present in the core communities of MBRAs of all different fecal types examined.

Additional file 6: Percent abundance of fecal OTUs absent from stable MBRA core communities. Table providing the percent abundance of abundant (>0.5 %) fecal OTUs absent from stable MBRA core communities.

Additional file 7: Bray-Curtis similarities between MBRA communities and their fecal samples determined from OTUs of different phyla. Table listing Bray-Curtis similarities between replicate MBRA communities of the same fecal type and between replicate MBRA communities and their starting fecal inocula based upon all OTUs as well as by OTUs partitioned by phyla.

Additional file 8: Analysis of abundance of core OTUs identified in fecal donor B MBRA communities as a function of time in culture. Heatmap presenting abundance of OTUs that were identified as present in individual core communities for MBRAs inoculated with donor B as well as abundant in the donor B fecal sample.

Additional file 9: Analysis of abundance of core OTUs identified in fecal donor C MBRA communities as a function of time in culture. Heatmap presenting abundance of OTUs that were identified as present in individual core communities for MBRAs inoculated with donor C as well as OTUs abundant in the donor C fecal sample.

Additional file 10: Comparison of reagent concentrations described in 1 L of medium. Table comparing reagent concentrations described for media used for cultivation of complex fecal communities from 11 different sources.

Additional file 11: Fresh and frozen fecal samples form similar communities in MBRA. NMDS plot and ANOSIM, PERMANOVA, and ANOVA of β-dispersion statistics of Bray-Curtis dissimilarities calculated from OTUs present in MBRA communities seeded with donor A2-fresh and donor A2-frozen fecal samples.

Additional file 12: Supplemental methods. Supplemental methods describing methods used to compare direct sequencing approach to DNA extraction and amplification as well as identification of potential sequencing contaminants from negative control samples.

Additional file 13: Impact of sample preparation method on Bray-Curtis and Sorenson similarities in replicate samples. Table providing differences in Bray-Curtis and Sorenson similarities in replicate samples prepared with different extraction methods.

Additional file 14: Distribution of sequences in shared OTUs across replicate samples prepared by different methods. A table providing percent differences in abundance of OTUs and sequences based upon sample preparation method and a graph plotting the abundance of different taxa across replicate samples prepared with different methods.

Additional file 15: Mean abundance across samples of 25 most abundant OTUs in negative control samples. Table listing the mean abundance across all samples sequences of 25 most abundant OTUs detected in negative control samples.

Abbreviations

ANI: average nucleotide identity; BC: Bray-Curtis; MBRA: minibioreactor array; OTU: operational taxonomic unit; SD: standard deviation.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

JMA contributed to the design of experiments, collected and analyzed the data, and drafted the manuscript. CDR contributed to the design of experiments and collected data. RAB contributed to design of experiments and drafting of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors acknowledge Quinn Barrett, Micah Keidan, Lilian Jensen, Ashley Thomas, and Bernard Schroeter for the technical assistance as well as two anonymous reviewers for helpful critiques on the manuscript. This work was supported by award 5U19AI090872-02 from the National Institutes of Allergy and Infectious Diseases to R.A.B.

Author details

[1]Center for Metagenomics and Microbiome Research, Department of Molecular Virology and Microbiology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA. [2]Department of Microbiology and Molecular Genetics, Michigan State University, 567 Wilson Rd, East Lansing, MI 48823, USA. [3]Present address: Institute for Molecular Biology, University of Oregon, 249 Klamath, Eugene, OR 97403, USA.

Received: 5 April 2015 Accepted: 1 September 2015
Published online: 30 September 2015

References

1. Blumberg R, Powrie F. Microbiota, disease, and back to health: a metastable journey. Sci Transl Med. 2012;4:137rv7.
2. Hooper LV, Littman DR, Macpherson AJ. Interactions between the microbiota and the immune system. Science. 2012;336:1268–73.
3. Nicholson JK, Holmes E, Kinross J, Burcelin R, Gibson G, Jia W, et al. Host-gut microbiota metabolic interactions. Science. 2012;336:1262–7.
4. Ursell LK, Treuren WV, Metcalf JL, Pirrung M, Gewirtz A, Knight R. Replenishing our defensive microbes. BioEssays. 2013;35:810–7.
5. Pflughoeft KJ, Versalovic J. Human microbiome in health and disease. Annu Rev Pathol Mech Dis. 2012;7:99–122.
6. Clemente JC, Ursell LK, Parfrey LW, Knight R. The impact of the gut microbiota on human health: an integrative view. Cell. 2012;148:1258–70.
7. Cho I, Blaser MJ. The human microbiome: at the interface of health and disease. Nat Rev Genet. 2012;13:260–70.
8. Guarner F, Malagelada J-R. Gut flora in health and disease. Lancet. 2003;361:512–9.
9. Wu GD, Chen J, Hoffmann C, Bittinger K, Chen YY, Keilbaugh SA, et al. Linking long-term dietary patterns with gut microbial enterotypes. Science. 2011;334:105–8.
10. Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermúdez-Humarán LG, Gratadoux J-J, et al. Faecalibacterium prausnitzii is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. Proc Natl Acad Sci U S A. 2008;105:16731–6.
11. Chang JY, Antonopoulos DA, Kalra A, Tonelli A, Khalife WT, Schmidt TM, et al. Decreased diversity of the fecal microbiome in recurrent Clostridium difficile-associated Diarrhea. J Infect Dis. 2008;197:435–8.
12. Gootenberg DB, Turnbaugh PJ. Humanized animal models of the microbiome. J Anim Sci. 2010;89:1531–7.
13. Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, DuGar B, et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. Nature. 2011;472:57–63.
14. Goodman AL, Kallstrom G, Faith JJ, Reyes A, Moore A, Dantas G, et al. Extensive personal human gut microbiota culture collections characterized and manipulated in gnotobiotic mice. Proc Natl Acad Sci U S A. 2011;108:6252–7.
15. Bäckhed F, Ding H, Wang T, Hooper LV, Koh GY, Nagy A, et al. The gut microbiota as an environmental factor that regulates fat storage. Proc Natl Acad Sci U S A. 2004;101:15718–23.
16. Nguyen TLA, Vieira-Silva S, Liston A, Raes J. How informative is the mouse for human gut microbiota research? Dis Model Mech. 2015;8:1–16.
17. Allison C, McFarlane C, MacFarlane GT. Studies on mixed populations of human intestinal bacteria grown in single-stage and multistage continuous culture systems. Appl Environ Microbiol. 1989;55:672–8.
18. Macfarlane GT, Macfarlane S, Gibson GR. Validation of a three-stage compound continuous culture system for investigating the effect of retention time on the ecology and metabolism of bacteria in the human colon. Microb Ecol. 1998;35:1–8.
19. Minekus M, Smeets-Peeters M, Bernalier A, Marol-Bonnin S, Havenaar R, Marteau P, et al. A computer-controlled system to simulate conditions of the large intestine with peristaltic mixing, water absorption and absorption of fermentation products. Appl Microbiol Biotechnol. 1999;53:108–14.
20. Freeman J, O'Neill FJ, Wilcox MH. Effects of cefotaxime and desacetylcefotaxime upon Clostridium difficile proliferation and toxin production in a triple-stage chemostat model of the human gut. J Antimicrob Chemother. 2003;52:96–102.

21. Molly K, Vande Woestyne M, Verstraete W. Development of a 5-step multi-chamber reactor as a simulation of the human intestinal microbial ecosystem. Appl Microbiol Biotechnol. 1993;39:254–8.
22. Honda H, Gibson GR, Farmer S, Keller D, McCartney AL. Use of a continuous culture fermentation system to investigate the effect of GanedenBC30 (Bacillus coagulans GBI-30, 6086) supplementation on pathogen survival in the human gut microbiota. Anaerobe. 2011;17:36–42.
23. Gibson GR, Fuller R. Aspects of in vitro and in vivo research approaches directed toward identifying probiotics and prebiotics for human use. J Nutr. 2000;130:391S–5.
24. Macfarlane GT, Macfarlane S. Models for intestinal fermentation: association between food components, delivery systems, bioavailability and functional interactions in the gut. Curr Opin Biotechnol. 2007;18:156–62.
25. Sousa T, Paterson R, Moore V, Carlsson A, Abrahamsson B, Basit AW. The gastrointestinal microbiota as a site for the biotransformation of drugs. Int J Pharm. 2008;363:1–25.
26. Molly K, Woestyne MV, Smet ID, Verstraete W. Validation of the simulator of the human intestinal microbial ecosystem (SHIME) reactor using microorganism-associated activities. Microb Ecol Health Dis. 1994;7:191–200.
27. Van den Abbeele P, Grootaert C, Marzorati M, Possemiers S, Verstraete W, Gérard P, et al. Microbial community development in a dynamic gut model is reproducible, colon region specific, and selective for Bacteroidetes and Clostridium cluster IX. Appl Environ Microbiol. 2010;76:5237–46.
28. Rajilic-Stojanovic M, Maathuis A, Heilig HGHJ, Venema K, de Vos WM, Smidt H. Evaluating the microbial diversity of an in vitro model of the human large intestine by phylogenetic microarray analysis. Microbiology. 2010;156(11):3270–81.
29. Robinson CD, Auchtung JM, Collins J, Britton RA. Epidemic Clostridium difficile strains demonstrate increased competitive fitness compared to nonepidemic isolates. Infect Immun. 2014;82(7):2815–25.
30. Legendre P, Legendre LF. Numerical Ecology. 3rd ed. Amsterdam: Elsevier; 2012.
31. Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. Nature. 2012;486:207–14.
32. Clarke KR. Non-parametric multivariate analyses of changes in community structure. Aust J Ecol. 1993;18:117–43.
33. Anderson MJ. A new method for non-parametric multivariate analysis of variance. Austral Ecol. 2001;26:32–46.
34. Werner JJ, Knights D, Garcia ML, Scalfone NB, Smith S, Yarasheski K, et al. Bacterial community structures are unique and resilient in full-scale bioenergy systems. Proc Natl Acad Sci U S A. 2011;108:4158–63.
35. Schloss PD, Schubert AM, Zackular JP, Iverson KD, Young VB, Petrosino JF. Stabilization of the murine gut microbiome following weaning. Gut Microbes. 2014;3:383–93.
36. Zhou J, Wu L, Deng Y, Zhi X, Jiang Y-H, Tu Q, et al. Reproducibility and quantitation of amplicon sequencing-based detection. ISME J. 2011;5:1303–13.
37. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol. 2013;79:5112–20.
38. Flores GE, Caporaso JG, Henley JB, Rideout JR, Domogala D, Chase J, et al. Temporal variability is a personalized feature of the human microbiome. Genome Biol. 2014;15:531.
39. Van den Abbeele P, Belzer C, Goossens M, Kleerebezem M, de Vos WM, Thas O, et al. Butyrate-producing Clostridium cluster XIVa speciesspecifically colonize mucins in an in vitro gut model. ISME J. 2012;7:949–61.
40. Kettle H, Louis P, Holtrop G, Duncan SH, Flint HJ. Modelling the emergent dynamics and major metabolites of the human colonic microbiota. Environ Microbiol. 2014;17:1615–30.
41. Nakamura S. Bioavailability of cellobiose and other non-digestible and/or non-absorbable sugar substitutes and related topics. Nutrition. 2005;21:1158–9.
42. Cherbut C. Inulin and oligofructose in the dietary fibre concept. Brit J Nutr. 2002;87:S159–62.
43. Terpend K, Possemiers S, Daguet D, Marzorati M. Arabinogalactan and fructo-oligosaccharides have a different fermentation profile in the Simulator of the Human Intestinal Microbial Ecosystem (SHIME). Environ Microbiol Rep. 2013;5:595–603.
44. Johnsson T, Nikkila P, Toivonen L, Rosenqvist H, Laakso S. Cellular fatty acid profiles of Lactobacillus and Lactococcus strains in relation to the oleic acid content of the cultivation medium. Appl Environ Microbiol. 1995;61:4497–9.
45. Li JY, Zhang LW, Du M, Han X, Yi HX, Guo CF, et al. Effect of tween series on growth and cis-9, trans-11 conjugated linoleic acid production of Lactobacillus acidophilus F0221 in the presence of bile salts. Int J Mol Sci. 2011;12:9138–54.
46. Derrien M, Collado MC, Ben-Amor K, Salminen S, de Vos WM. The mucin degrader Akkermansia muciniphila is an abundant resident of the human intestinal tract. Appl Environ Microbiol. 2008;74:1646–8.
47. Marzorati M, Vanhoecke B, De Ryck T, Sadabad MS, Pinheiro I, Possemiers S, et al. The HMITM module: a new tool to study the host-microbiota interaction in the humangastrointestinal tract in vitro. BMC Microbiol. 2014;14:133.
48. Ulluwishewa D, Anderson RC, Young W, McNabb WC, van Baarlen P, Moughan PJ, et al. Live Faecalibacterium prausnitzii in an apical anaerobic model of the intestinal epithelial barrier. Cell Microbiol. 2015;17:226–40.
49. Stelzner M, Helmrath M, Dunn JCY, Henning SJ, Houchen CW, Kuo C, et al. A nomenclature for intestinal in vitro cultures. Am J Physiol Gastrointest Liver Physiol. 2012;302:G1359–63.
50. Kovbasnjuk O, Zachos NC, In J, Foulke-Abel J, Ettayebi K, Hyser JM, et al. Human enteroids: preclinical models of non-inflammatory diarrhea. Stem Cell Res Ther. 2013;4:S3.
51. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J. 2012;6:1621–4.
52. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, *et al.* vegan: Community Ecology Package. R package version 2.0-9. http://CRAN.R-project.org/package=vegan.
53. Rajaram S, Oono Y. NeatMap-non-clustering heat map alternatives in R. BMC Bioinformatics. 2010;11:45.

DRAFT

STRATEGIC PLAN FOR NIH NUTRITION RESEARCH

DRAFT

TABLE OF CONTENTS

EXECUTIVE SUMMARY 5

INTRODUCTION AND BACKGROUND 6

BACKGROUND AND SIGNIFICANCE 6

CROSS-CUTTING RESEARCH AREAS 10

Minority Health and Health Disparities 10

Women's Health 10

Rigor and Reproducibility 11

Systems Science 12

PROCESS FOR DEVELOPING THE STRATEGIC PLAN FOR NIH NUTRITION RESEARCH 13

STRATEGIC PLAN THEMES 15

THEME 1: INVESTIGATE NUTRITIONAL BIOCHEMISTRY, PHYSIOLOGY, AND THE MICROBIOME 15

INTRODUCTION 15

SCIENTIFIC PRIORITIES 16

1-1. Advance Nutritional Biochemistry and the Bioinformatics of Nutrition-Related Pathways 16

1-2. Clarify the Integrative Physiology of Ingestive Behaviors 17

1-3. Identify and Leverage Interrelationships between Diet, Host, and the Gut Microbiome to Promote Health 18

THEME 2: ASSESS THE ROLE OF NUTRITION AND DIETARY PATTERNS IN DEVELOPMENT, HEALTH, AND DISEASE ACROSS LIFE STAGES 23

INTRODUCTION 23

SCIENTIFIC PRIORITIES 24

2-1. Examine the Role of Prenatal Nutrition for the Health of Mother and Offspring 24

2-2. Research Nutrition in the Birth-to-24 Month Period 25

2-3. Optimize Nutrition for Children and Adolescents 27

2-4. Determine Nutritional Requirements for Healthy Aging in Older Adults 28

2-5. Determine Mechanisms by which Dietary Patterns Affect Health Status and Chronic Disease Susceptibility 29

THEME 3. EXPLORE INDIVIDUAL VARIABILITY IN RESPONSE TO DIET INTERVENTIONS TO INFORM NUTRITION SCIENCE, IMPROVE HEALTH, AND PREVENT DISEASE 33

INTRODUCTION 33

SCIENTIFIC PRIORITIES 34

3-1. Elucidate the Biological Factors Underlying Individual Variation in Response to Dietary Interventions . 34

3-2. Examine the Contribution of Social Determinants of Health and Other Environmental Factors to Variations in Response to Nutrition Interventions 35

THEME 4: ENHANCE CLINICAL NUTRITION RESEARCH TO IMPROVE HEALTH OUTCOMES IN PATIENTS 37

INTRODUCTION 37

SCIENTIFIC PRIORITIES 38

4-1. Identify and Leverage Interactions Among Nutrition, Disease States, and Treatments 38

4-2. Improve Assessment of Malnutrition 39

4-3. Identify Triggers and Endpoints for Nutritional Support in Clinical Settings 40

4-4. Optimize Nutritional Support for Premature and Low Birthweight Infants 41

4-5. Optimize Nutrition Interventions for Patients with Special Nutritional Requirements 42

THEME 5: ADVANCE IMPLEMENTATION SCIENCE TO INCREASE THE USE OF EFFECTIVE NUTRITION INTERVENTIONS ... 43

INTRODUCTION 43

SCIENTIFIC PRIORITIES 44

5-1. Evaluate Strategies for Sustaining Nutrition-Related Behavior Change 44

5-2. Leverage Multidimensional Strategies to Increase Use of Evidence-Based Nutrition Approaches 45

THEME 6: DEVELOP AND REFINE RESEARCH METHODS AND TOOLS 47

INTRODUCTION 47

SCIENTIFIC PRIORITIES 47

6-1. Improve Methods to Assess Dietary Intake and Exposure 47

6-2 Develop and Validate Biomarkers of Relevance for Chronic Disease and Diet 51

6-3. Improve Methods for Dietary Pattern Analysis 52

6-4. Develop and Improve Tools Using Big Data for Systems Science Approaches to Nutrition Research 53

6-5. Develop Sensors for Continuous Monitoring of Nutrients and Metabolites for Personalized Nutrition Research 54

6-6. Develop Predictive Epigenetic Tools 55

6-7. Encourage the Use of Controlled Human Feeding Studies 56

THEME 7: SUPPORT TRAINING TO BUILD AN OUTSTANDING NUTRITION RESEARCH WORKFORCE 57

INTRODUCTION 57

TRAINING PRIORITIES 57

7-1. Facilitate Training in Host, Gut Microbiome Metabolism, and Diet Interrelationships 57

7-2. Enhance Training in the Application of Big Data Approaches to Nutrition Research 58

CONCLUSION 60

GLOSSARY 61

APPENDIX A: MEMBERSHIP OF THE NIH NUTRITION RESEARCH TASK FORCE 65
NRTF SENIOR LEADERSHIP GROUP 65
NRTF WORKING GROUP 66
APPENDIX B: MEMBERSHIP OF THE STRATEGIC PLAN THOUGHT LEADER PANEL 69
APPENDIX C: ACKNOWLEDGMENTS 71

EXECUTIVE SUMMARY

PLACE HOLDER – THIS SECTION WILL BE ADDED AFTER THE PUBLIC COMMENT PERIOD

INTRODUCTION AND BACKGROUND

Background and Significance

Nutrition encompasses the processes involved in ingesting food and dietary components and their absorption and utilization for growth and development, metabolism, repair, and health maintenance. Nutrition is integral to the prevention and treatment of disease as well as for health promotion. Poor nutrition contributes to some of the leading causes of death and increases the risk of numerous diseases, including heart disease, diabetes, obesity, high blood pressure, stroke, cancer, and osteoporosis. Improving nutrition has the potential to be one of the most cost-effective health care strategies for reducing human morbidity and mortality across the world.

The National Institutes of Health (NIH) leads all federal agencies in funding nutrition research and research training. The NIH funds extramural research at institutions throughout the country, as well as intramural research at the NIH. For decades, the NIH has supported basic and clinical research, epidemiologic research, translational/application-based studies, and research training in a range of nutrition-related areas across the lifespan. Because diet and nutritional status are among the most influential determinants of health, nutrition research is supported by many Institutes, Centers, and Offices (ICOs) of the NIH.

NIH-funded research has led to some of the most important advances in nutrition research to date. For example, studies have identified specific nutrient-disease relationships, such as the relationship between dietary saturated or *trans* fatty acids and blood low-density lipoprotein (LDL) cholesterol. Landmark NIH-supported trials have successfully shown the impact of dietary patterns or nutrient combinations on disease prevention or treatment. For example, Dietary Approaches to Stop Hypertension (DASH) a multicenter, randomized, controlled-feeding trial, revealed how dietary approaches can be used to prevent and treat hypertension. The second Age-Related Eye Disease Study (AREDS2) determined that a nutritional supplement containing vitamin C, vitamin E, zinc, copper, lutein, and zeaxanthin can reduce the rate of deterioration in age-related macular degeneration. The Diabetes Prevention Program (DPP) clinical trial demonstrated that diet and exercise changes designed to achieve 7 percent weight loss can prevent or delay the onset of type 2 diabetes; reduce the need for antidiabetic, lipid-lowering, and antihypertensive medications; and improve overall health in people at risk of developing type 2 diabetes. NIH-funded research has also enabled substantial advances in the understanding of the interactions between the diet, environment, and microbiome, and has

facilitated continual refinement of dietary assessment methodology through improved data capture and analytical approaches.

Research funded by the NIH generates much of the scientific evidence used to establish the Dietary Reference Intakes (DRIs), a set of reference values used to guide and assess nutrient intakes of healthy people, and to develop the *Dietary Guidelines for Americans*, a set of recommendations on the components of a healthy and nutritionally adequate diet. Other federal agencies, such as the U.S. Food and Drug Administration (FDA), rely on research supported by the NIH to make evidence-based regulatory and policy decisions. In sum, NIH nutrition research has the potential to have a broad impact on public health.

Although much has been learned to date about nutrition and health, gaps in knowledge remain. The relationships between nutrition and health are extremely complex, requiring cross-disciplinary expertise to better explain these connections. The relationships between nutrients and biological systems are reciprocal and dynamic. Recent technologic advances enhance the potential for further discovery in this area. In addition, emerging evidence of the impact of psychosocial, sociocultural, and environmental factors on these relationships represents an opportunity to identify more effective approaches to addressing nutrition needs across an increasingly diverse U.S. population. Building upon NIH's investment in nutrition research to address known research gaps and pursue new scientific opportunities in the next decade will require collaborative efforts across the NIH.

To make the best use of NIH's collective knowledge and resources to inform decisions most effectively over the next decade, Dr. Francis Collins, the Director of NIH, established the NIH Nutrition Research Task Force (NRTF) in October 2016 to coordinate and accelerate progress in NIH-funded nutrition research and guide the development of the first 10-year NIH-wide strategic plan for nutrition research.

This document, the *Strategic Plan for NIH Nutrition Research*, will serve as a guide to accelerate basic, translational, and clinical research, as well as research training activities, over the next 10 years. The opportunities to address the priorities identified in this plan are, in many cases, made possible by advances in bioinformatics (i.e., the integration of computer hardware and software tools, imaging technology, and databases to address biological questions), computational modelling, and discovery technologies (e.g., metabolomics, genomics, transcriptomics, proteomics) that can be applied to help answer many nutrition-related questions.

The charge to develop the *Strategic Plan for NIH Nutrition Research* dovetails with the 2016 *National Nutrition Research Roadmap*, created by the Interagency Committee on Human Nutrition Research, a trans-federal government committee charged with improving the coordination and communication among multiple federal agencies engaged in nutrition research.[1] The *Strategic Plan for NIH Nutrition Research* was developed with the broad research goals of the *National Nutrition Research Roadmap* in mind, but it focuses specifically on priorities best aligned with the mission of the NIH, which is to seek fundamental knowledge about the nature and behavior of living systems and to apply that knowledge to enhance health, lengthen life, and reduce illness and disability.

The *Strategic Plan for NIH Nutrition Research* is complementary to related disease-specific research plans previously developed by the NIH, such as the *Strategic Plan for NIH Obesity Research.*[2] However, the *Strategic Plan for NIH Nutrition Research* specifically focuses on the impact of diet and nutrition related to all health conditions. The *Strategic Plan* also aligns with recommendations of the *NIH-Wide Strategic Plan, Fiscal Years 2016–2020: Turning Discovery Into Health,*[3] which outlines a vision for biomedical research to capitalize on new opportunities for scientific exploration and address new challenges for human health.

Because the relationships between food, nutrition, and health are so complex, multifaceted research priorities are outlined around the following seven Themes and cross-cutting research areas. Each Theme contains major research priorities and examples of related research activities. The order of these Themes, and the number of identified research priorities and activities, are not indicative of any relative emphasis by the NIH:

1. Investigate Nutritional Biochemistry, Physiology, and the Microbiome
2. Assess the Role of Nutrition and Dietary Patterns in Development, Health, and Disease Across Life Stages
3. Explore Individual Variability in Response to Diet Interventions to Inform Nutrition Science, Improve Health, and Prevent Disease

[1] https://www.nal.usda.gov/fnic/interagency-committee-human-nutrition-research
[2] https://obesityresearch.nih.gov/about/strategic-plan.aspx
[3] https://www.nih.gov/about-nih/nih-wide-strategic-plan

4. Enhance Clinical Nutrition Research to Improve Health Outcomes in Patients
5. Advance Implementation Science to Increase the Use of Effective Nutritional Interventions
6. Develop and Refine Research Methods and Tools
7. Support Training to Build an Outstanding Nutrition Research Workforce

Each Theme encompasses both new areas where the NIH is strategically positioned to make an impact and existing areas of research that hold considerable potential for growth and expansion over the next 10 years. Due to the integrated nature of the seven Themes, it is expected that advances in any one of the thematic areas will positively influence the potential for advances in other areas. Although the Task Force identified seven Themes and multiple priority research areas and activities within each Theme, it is recognized that opportunities exist for NIH-funded researchers to also address issues beyond these thematic areas. Also, as with any future-oriented plan, the Strategic Plan will evolve as research progresses and new opportunities and challenges emerge.

Cross-Cutting Research Areas

Several cross-cutting research areas are relevant across all Themes in the *Strategic Plan for NIH Nutrition Research*. These areas were identified by the Task Force as critical for nutrition research, but they also are reflected in other NIH strategic plans as important considerations in scientific research more broadly. Cross-cutting research areas include Minority Health and Health Disparities, Women's Health, Rigor and Reproducibility, and Systems Science.

Minority Health and Health Disparities

Many populations, whether defined by race, ethnicity, socioeconomic status, disability, sex, gender, or geography experience higher rates of certain diseases and greater mortality, when compared with the general population. Nutrition-related health disparities arise from multiple factors operating within biological, behavioral, and environmental domains and across individual, interpersonal, community, and societal levels of influence.

For many chronic diseases and conditions, diet is a key contributor to health disparity. Because factors that contribute to diet-related health disparities are complex, they may individually or synergistically affect interrelationships among dietary intake, nutritional status, and health. Obtaining a better understanding of these interrelationships can help to elucidate how disparities in diet and nutrition lead to adverse health consequences and has potential implications for public policy. Issues related to health disparities are therefore integral to all Themes of the *Strategic Plan*.

Women's Health

Nutrition plays an important role in many diseases and conditions that primarily affect women or affect woman differently than men. One goal related to this cross-cutting research area will be to expand research on the impact of nutrition and diet-related disorders during pregnancy and the effects on women's health and risk of future disease. Two priorities in this area are inclusion of women in clinical research and consideration of sex as a biological variable in basic and preclinical research related to nutrition. The *Strategic Plan* aims to stimulate and encourage basic, translational, and clinical nutrition research on the role of sex and gender in health and disease across all the Themes.

Rigor and Reproducibility

The NIH is committed to promoting rigorous and reproducible research along with impartial and transparent communication of that research in all areas of biomedical science, including nutrition. Thus, the NIH-wide strategic goal to increase rigor and reproducibility in biomedical research is a cross-cutting research area addressed by this *Strategic Plan.* Rigor and reproducibility should be considered in the design, conduct, analysis, and communication of nutrition research. To this end, opportunities to establish best practices during all phases of research and to develop more rigorous research methodologies and analytic strategies are described throughout the plan.

Harmonization. One avenue to advance scientific rigor and increase reproducibility is to standardize and harmonize nutrition research methods, measures, and data elements. This approach would enable investigators to more easily compare or pool results across studies, thus reinforcing the strength of findings and decreasing research costs. Improving dietary assessment methods, decreasing errors associated with assessment methods, and developing approaches to incorporate multiple methods of dietary assessment into research for cross-validation are additional opportunities to improve the rigor and reproducibility of nutrition research.

Nutrition as a Relevant Variable Influencing all Biological Responses. One goal of the NIH's efforts to improve rigor and reproducibility in biomedical research is to better understand the relevant factors or variables that can contribute to situations where similarly designed studies give rise to different outcomes or conclusions. One of the recent updates to NIH grant proposal and review instructions asks investigators to consider relevant biological variables including sex, age, weight, and underlying health conditions during the planning, implementation, analysis, and communication of biomedical research. The underlying rationale is that these are often critical factors affecting health or disease that, if ignored, may lead to incomplete understanding of biological functions, disease processes, and treatment responses.

Given the well-known causal links between diet, nutritional status, growth, development, health, and disease, it is reasonable to hypothesize that differences in dietary intake and nutritional status may also be a significant source of variability in biomedical research and may contribute to a lack of reproducibility between studies. Thus, this *Strategic Plan* proposes that dietary intake data and markers of nutritional status be collected more broadly across a wide range of biomedical research studies, particularly for those in which some evidence exists to suggest that differences in dietary intake and nutritional status may alter responses to

interventions. In addition, including nutrition scientists in the design, analysis, and communication of research results may help to improve reproducibility.

Rigor and Transparency in Conducting Nutrition Research. Efforts to improve rigor and transparency should be considered in the planning stages and throughout implementation, analysis, and communication of research. Opportunities to improve transparency can be found throughout the *Strategic Plan*. For example, in preclinical research, what diet should be used is frequently not a consideration and is sometimes not even reported. When publishing research, transparency about the details of the diets and the rationale for choosing them should be encouraged as this could be a significant source of unexplained variability and lack of reproducibility.

Systems Science

Foods contain multiple nutrients and bioactive components that interact synergistically or antagonistically with many different systems (e.g., biological, behavioral, social, and environmental) in multiple and complex ways. A variety of methodological approaches could be employed to study these relationships. Systems science approaches, including computational modeling, may allow researchers to better understand the dynamic and adaptive processes involved in nutrition, personalized nutrition, and gene-environment interactions. For certain questions, systems science may be uniquely positioned to help explain how a variety of factors interact to achieve a specific nutrition-related outcome (e.g., elucidating the complex biological systems, showing how these systems interact with other social and behavioral systems, and connecting nutrition with disease and other aspects of human health). Systems science is useful not only for investigating complex diet-disease relationships but also for designing and evaluating interventions and policies that address socioeconomic and cultural issues underpinning unhealthy dietary behaviors. Using a systems science approach to study interactions among diet, biological systems, and health status will be imperative to advancing the nutrition research agenda at the NIH.

Process for Developing the Strategic Plan for NIH Nutrition Research

To coordinate and accelerate progress in nutrition research, Dr. Francis Collins, NIH Director, established the Nutrition Research Task Force in 2016 to develop and implement this first 10-year *Strategic Plan for NIH Nutrition Research*. The Task Force is chaired by Dr. Griffin P. Rodgers, Director of the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK). Co-chairs are Dr. Gary Gibbons, Director of the National Heart, Lung, and Blood Institute; Dr. Norman Sharpless, Director of the National Cancer Institute; and Dr. Diana W. Bianchi, Director of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development. Dr. Christopher Lynch, Director of NIDDK's Office of Nutrition Research, serves as the Task Force's Executive Secretary. To accomplish its mission, the Task Force formed a trans-NIH Nutrition Senior Leadership Group along with a Working Group drawn from many ICOs within NIH (see Appendix A for a full membership list).

To obtain information about research gaps and opportunities that should be addressed by this *Strategic Plan*, the Task Force sought the assistance of the external nutrition research community through several mechanisms. Members of the Task Force conducted a scan of recent literature to identify research gaps that had been captured by other federal agencies, professional societies, and researchers working across the field of nutrition. More than 50 relevant publications were identified. The Task Force then used an online crowdsourcing platform to solicit information from the public regarding the critical gaps and opportunities in nutrition research that should be addressed by the NIH over the next 10 years. Invitations encouraging participation in the crowdsourcing effort were sent to more than 40 professional societies and thousands of academic researchers. A link to the crowdsourcing webpage was also available on the Office of Nutrition Research website. The nutrition-related research gaps and opportunities gleaned from the literature search and crowdsourcing efforts were combined into one comprehensive document that was shared with a Thought Leader Panel of 30 external experts from across diverse areas of nutrition science (see Appendix B for a list of panelists). These experts were invited to provide input on prioritizing these research opportunities.

After reviewing the collective suggestions from the Thought Leader Panel and broader nutrition research community, members of the Task Force identified seven Themes that reflected the priorities they felt were most important to encourage among both the extramural community and the NIH intramural research program. These Themes and corresponding Priorities are described in detail in this *Strategic Plan*.

Public Comments: Place holder, this section will be written after public comment period

STRATEGIC PLAN THEMES

Theme 1: Investigate Nutritional Biochemistry, Physiology, and the Microbiome

Introduction

Major gaps remain in the current understanding of the genes, mechanisms, and pathways involved in the physiological and behavioral responses to diet. Scientists are still learning about processes that play a role in food selection and consumption, including specific type, timing, and amount of food consumed. Thus, a priority for Theme 1 is to elucidate the drivers and integrative pathways involved in ingestive behaviors, which encompass all eating and drinking behaviors and the associated biological, psychological, and sociocultural modifiers. For example, research to elucidate the peripheral actions and central neural circuitry that are altered by the presence of nutrients or changes in nutrient status may help to explain certain ingestive behaviors. Research is needed to more fully explain the pathways that mediate hedonic food responses and food choice. How changes in nutrient and health status, life stage, and chronobiology are interrelated with ingestion is also a gap that could be addressed by deeper exploration of these relationships.

Another emerging area of interest and opportunity involves learning more about the bidirectional interrelationships between diet, the activities of microbes that live within the body (the host), and the host's nutritional status, health, disease susceptibility, and ingestive behaviors and responses. Systems biology, an approach which aims to describe the overall behavior of a biological system through quantitative experimentation and computational modeling, is one way to explore these complicated interrelationships. The foundational biochemistry, physiology, and microbial data emerging from this Theme will help inform the bioinformatics of nutrition-related genes and processes that will help accelerate the use of systems biology and systems science approaches for nutrition research.

Scientific Priorities

1-1. Advance Nutritional Biochemistry and the Bioinformatics of Nutrition-Related Pathways

Bioinformatic gaps exist for many of the genes whose proteins are responsible for digestion, absorption, distribution, metabolism, excretion, and storage of nutrients throughout the body and for many of the genes engaged in the monitoring, use of, and responses to the presence of these nutrients. The following examples illustrate some of the many areas in which information is lacking or where further progress is needed.

The first example encompasses the genes involved in the transport of nutrients and metabolites across cell and organelle membranes and the need to improve knowledge in this area. Characterization of many of the transporter genes regarding solute specificity, tissue-specific expression, subcellular localization, allelic variation, and alternate splicing and regulation could be improved. This is also true for some tastant and odorant receptors. These receptors and their associated neural pathways serve as the gateway for regulation of food intake and dietary choices, and thereby have a considerable potential to affect health. A third example that could benefit from additional research relates to other nutrient and metabolite monitoring and response systems throughout the body. In these systems, the genes and pathways involved in intracellular signaling mechanisms need to be better characterized.

Further characterization of nutrition-related genes and their products, along with their allelic and splice variants, may help explain some inter-individual variability in responses to dietary interventions. Research on the regulation of taste and olfactory pathways will help explain the role of such regulation in hedonic responses and taste preferences. Overall, this research is needed as a first step toward refining systems biology approaches to nutrition research.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Characterize the nutrient specificity, physiological roles, regulation, allelic variation, tissue specificity, subcellular localization, and age-related roles (e.g., fetal versus adult) of nutrition-relevant genes and how they are dysregulated during disease processes.
- Characterize the genes, genetic modifications, and cell signaling mechanisms of systems involved in monitoring and responding to nutrient or metabolite status or intake and their roles in health and disease.

- Elucidate the molecular targets and mechanisms involved in nutrient and metabolite sensing and determine how their relevant signaling is integrated with or affected by other factors, including disease conditions.
- Determine the cellular mechanisms involved in the oral sensing of food constituents and the regulation of taste receptors and olfactory receptor cell signaling involved in nutrient intake and sensing.

1-2. Clarify the Integrative Physiology of Ingestive Behaviors

This Priority focuses on how nutrient information is transmitted throughout the body, and the integration of information to regulate metabolism, physiology, and behavior. Integrative physiological pathways responsible for ingestive behaviors, including their associated biological, psychological, and social processes, need to be clarified. Examples of these pathways include gut-brain signaling by incretins, bile acids, vagal-vagal or gastrointestinal (GI) hormones, central nervous system (CNS) monitoring of nutrient status, and the mechanisms involved in translating the signaling into behavioral, metabolic, or physiological responses. Work is also needed to understand how external, internal, and circadian cues, and their interactions, affect these pathways.

Research is needed to clarify the functional effects of and integrative pathways responsible for nutrition-related GI peptides, hormones, and vagal activation, as well as adipose-derived lipokines, adipokines, cytokines, and small nucleic acids. This research could take advantage of new neuromodulatory technologies, including vagal nerve stimulation and blocking, in combination with imaging and blood sampling to elucidate mechanisms and pathways involved in normal and disrupted eating behaviors.

Research is also needed to determine the biological basis for food desire and food choice with the ultimate goal of developing new approaches to modify diet-related behaviors. This includes research at the neurophysiological level, and at broader levels in which the environment and behavior are involved. Such studies could take advantage of recent advances in conceptual and other behavioral frameworks, as well as genetic and molecular tools that can be applied in human and animal models to study the pathways, mechanisms, and substrates responsible for nutrition-related behaviors and food choice.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Identify the biological and environmental (e.g., early life feeding practices) variables that affect food choices for health promotion and disease prevention.
- Elucidate how feedback from the periphery alters CNS control of food intake.
- Establish the role of oral and extraoral (expressed throughout the body) taste and olfactory receptors in behavioral, hedonic, metabolic, or physiological responses.
- Identify factors that regulate or modulate hedonic systems relevant to ingestive behavior.
- Determine and leverage mechanisms whereby whole-body energy balance adapts to caloric excess or deficiency.

1-3. Identify and Leverage Interrelationships between Diet, Host, and the Gut Microbiome to Promote Health

Individual gut microbiota (including bacteria, fungi, Archea, bacteriophages, and viruses) have sometimes been designated as either beneficial or pathogenic. However, the gut microbiota increasingly has been recognized as an ecosystem based on symbiotic relationships (mutualism, commensalism, and parasitism) between members of the microbial communities and between these communities and the host. Recent research has revealed evidence of metabolic niches and microbial community metabolism in which individual species are responsible for different metabolic steps. Research is needed to clarify these niches, their community metabolism, and how diet may affect them.

Investments to improve methods and resources have led to a tremendous leap forward in the ability to pursue research on the gut microbiome (the collective genomes of the microbes that live inside and on the human body). This increased accessibility has fueled a groundswell of preclinical and clinical studies exploring potential roles of the gut microbiome in integrative physiology, health, behavior, and a broad range of chronic diseases.

Despite the recognition of the potential for food to affect microbiota metabolism in the gut, and to a lesser extent the species diversity, only a fraction of microbiome research has collected data on the diet or nutritional status of the host. Researchers have a tremendous opportunity to better understand how dietary components or dietary patterns affect the host through the gut microbiome, to elucidate the underlying mechanisms and pathways—including the gut microbiome-brain axis—and to leverage these interrelationships to promote the health of the

host. Facilitating improvements in research practices related to the inclusion of diet and nutrition considerations in microbiome research is a priority.

1-3a. Identify Best Practices in the Design of Research Examining Interrelationships Between the Host, Gut Microbiome, and Diet

The reproducibility of microbiome research continues to improve through efforts initiated and led by many groups. However, not yet addressed are issues related to the impact of host diet and nutritional and digestive health status on microbiome measures. Research opportunities to improve rigor and reproducibility include improving current practices or developing new best practices in the study design, analysis, and scientific reporting phases of microbiome research. An approach to facilitate achieving this goal is to encourage the inclusion of diet and nutritional status measures and nutrition expertise in microbiome research design. This is particularly important as it may improve the reproducibility of these studies.

Another opportunity to improve rigor and reproducibility pertains to both preclinical and clinical microbiome research, where careful consideration should be given to the selection of diet. The rationale for the choice of diet should be factored into the analysis and carefully reported in scientific communications. Attention should be focused on the type of dietary fiber (soluble versus insoluble) and whether it is metabolizable (accessible) by the microbiota or not. The chemical diversity of dietary fibers may present opportunities for selectively engaging microbiota-mediated metabolic pathways and should be taken into consideration when designing an experimental diet. The micronutrient and polyphenol content of the diet should also be considered, as these will affect microbial ecology, including microbiota composition and metabolic activities. Additionally, dietary manipulations to reflect the various types of diets consumed by different populations could provide information on gut microbial flora associated with these dietary patterns. The timing between diet interventions and microbiome sampling should also be considered and should be based on known transit times in the GI region of interest.

The following information is frequently obtained as part of clinical diagnostic investigations for digestive disorders: 1) the total number of microbes in a fecal sample, 2) stool quality (e.g., as measured by the Bristol Stool Scale), and 3) bowel habits. However, this information is not consistently collected in microbiome research studies. Collecting these standard measures of digestive health is a third area that could benefit gut microbiome research.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Prospectively investigate the role of the microbiome in the development of diet-related health outcomes.
- Elucidate the optimal range of times between the start of a dietary intervention and assessments to examine the consequences of that intervention on gut microbial ecology.
- Elucidate the bidirectional interactions between the host's diet, genome, metabolome, and gut microbiome and determine the diet-related factors that lead to the development of a stable microbiome during childhood.

1-3b. Encourage Development of Bioinformatic Resources and Tools

Bioinformatic approaches are often used when large datasets are generated, such as in microbiome research. As the understanding of the microbiome evolves, new tools and resources will be needed to address and advance microbiome-diet research.

Several approaches could be explored to foster research in this area. Improving the annotation of metabolomic and proteomic libraries is one of them. The knowledge of metabolites produced by the gut microbiota is very limited, and it is not known which of those molecules are derived from dietary sources. Improving this knowledge and developing a database of spectral features from fresh and prepared common foods (and food-based metabolites) along with partially digested or gut microbiota metabolized food is urgently needed. These data should be incorporated into existing databases.

A second important opportunity involves studies to elucidate microbial-mediated metabolic pathways. Further research is needed to understand the ecological flow of dietary substrates and their microbial metabolism in the gut, especially for microbial metabolites that affect the host. Diet serves as a substrate to produce metabolites that may have a significant effect on host physiology in health and disease and can affect the species composition of the gut microbiome.

Permitting real-time monitoring of GI metabolites along with *in situ* sampling of, and delivery of metabolites to, discrete regions of the GI tract is a third potentially fruitful avenue for research in this area. Nutrient and metabolite signaling in discrete regions of the GI tract may be important for tastant, incretin, GI reflexes (e.g., through gut-brain axis or enteroendocrine

responses), and inflammation of the gastrointestinal-associated lymphoid tissues. Many discrete gut regions, such as the jejunum, ileum, and cecum, are presently inaccessible by endoscopic or other approaches in conscious, fed individuals. New tools and technologies that would allow study of the microbiome *in situ* and in discrete regions of the gut could advance the field and better clarify the regional differences and dynamic nature of the responses of the microbes to diet. Moreover, such sampling could facilitate studies comparing mucosal versus luminally associated microbes.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Determine the molecular identity of unknown mass spectrometric signatures associated with microbial metabolism of ingested foods.
- Elucidate metabolic pathways that require enzymatic activities of multiple microorganisms, especially for metabolites of physiological or pathological significance for the host.
- Identify microbiota and metabolites in chyme from distinct regions of the gut after meal challenge or diet interventions.
- Investigate host-gut microbiota co-metabolites such as bile acids.

1-3c. Discover Mediators, Mechanisms, and Translational Roles of Diet-Host-Gut Microbiota Interrelationships

Gut microbes not only influence what the human host is able to extract from the diet, both nutritionally and energetically, but also may affect host physiology, behavior, and susceptibility to diet-related chronic disease. Research into the associations between diet and the gut microbiome is needed to further elucidate their bidirectional interactions, and how prebiotics, probiotics, and synbiotics influence these associations.

One goal of this Theme is to expand knowledge about how diet-gut microbiome interactions affect host physiology and behavior. Identifying the mechanisms underlying microbiome-mediated effects on health and disease susceptibility could lead to new therapeutic targets. Another goal is to identify roles of the regional GI microbial ecologies in host biology and health as well as their physical importance. Microorganisms in other gut regions, which may be closer to major axes of nutrient signaling and regulation would be of interest, as they may be more dynamic or responsive to diet than those of the cecum.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Develop methods to produce stable isotope-labeled dietary fibers and oligosaccharides that are digestible by the gut microbiota but not directly by the host.
- Determine the mechanisms by which diet influences the oral or gut microbial ecosystems to bring about changes in host biology, behavior, and health.
- Determine whether bioactive microbial metabolites, have a physiologic impact on human mucosal immunity and cardiovascular disease risk.
- Elucidate the roles of regional gut microbiota from oral cavity to colon and identify diet-dependent variability in these regional ecosystems and the consequences for the host.
- Investigate variations in microbial composition as a function of age and in the elderly.

Theme 2: Assess the Role of Nutrition and Dietary Patterns in Development, Health, and Disease across Life Stages

Introduction

It is known that nutritional requirements change across the lifespan. However, gaps exist in knowledge about the nutrient requirements for optimal health, particularly during the life stages of *in utero,* infancy, childhood, adolescence, young adulthood, and during pregnancy and lactation. Mapping the trajectory of nutrition and health relationships over the lifespan is a goal identified in the *Strategic Plan.* Research in this area will require the use of new technologies and systems biology approaches. Identifying appropriate normative data will lead to a clearer understanding of how nutrient requirements change and how best to intervene to support development, maturation, and aging. Determining critical time periods for intervention will provide useful targets for future recommendations.

Several life stages have been understudied and need immediate attention, including pregnancy and from birth to 24 months. It is understood that optimal nutrition during pregnancy is essential for the health of the mother and for fetal development, but recent research suggests that both prenatal and early life nutritional exposures may have lifelong consequences in the offspring's immune function, chronic disease risk, taste/food preferences, mental health, and cognition. Risk for metabolic and chronic diseases, including obesity, hypertension, and diabetes, may be programmed by nutritional status during the prenatal period. Animal models and limited human studies suggest that these effects may even be transgenerational through genomic imprinting. While a majority of NIH-supported research is conducted in the U.S., a limited amount of research that leverages special opportunities takes place elsewhere, and unique opportunities presented by past international studies have advanced knowledge of the role of nutrition in early development and other aspects of health. Additional research into the impact of early nutrition in the prenatal period and the first 24 months of life (i.e., the first 1,000 days), would inform development of dietary recommendations for this life stage. The first 1,000 days represent the largest gap and opportunity to advance knowledge of nutritional needs in children. However, nutrition needs in children and adolescents (2-18 years) and in young adults (19-35 years) also require more study.

Research at the other end of the age continuum is important and is needed to help determine optimal nutrition to support healthy aging. The proportion of the U.S. population older than age 65 years is projected to increase during the period covered by this *Strategic Plan* and

the population older than 65 is estimated to be about 78 million by 2035, according to the U.S. Census Bureau's 2017 National Population Projections.[4] The number of individuals older than 80 years is expected to rise to about 19 million during the same period. It should be anticipated that, with increased life expectancy and a growing number of older Americans, diet-related chronic diseases will become more prevalent, affecting older adults disproportionately. In addition, disabilities (e.g., frailty and cognitive declines) that inhibit the ability to live independently will increase as the population ages, and will exert demands on the public health, medical, and social service systems. Addressing malnutrition, as well as sarcopenia (loss of muscle tissue), in older adults is critical.

Scientific Priorities

2-1. Examine the Role of Prenatal Nutrition for the Health of Mother and Offspring

Despite the important associations between diet and health, nutrition during pregnancy and the role of maternal nutrition for fetal development remains understudied. This Priority has two goals. The first is to address questions around the optimal dietary requirements and/or patterns to avoid deficiencies, support maternal and fetal health, and set the stage for a healthy life course. Relationships between nutritional status and developmental and long-term health outcomes are complex, as they may be influenced by many biological factors (e.g., stress, inflammation, hormones, genetics, body composition) as well as by lifestyle behaviors, health disparities, and sociocultural factors. More research is needed to fully explain these and other unknown factors. The impact of nutrition during pregnancy on the mother and the fetus also varies depending on the stage of fetal development. Thus, it will be important to consider development when designing and evaluating interventions and interpreting study results.

The second goal is to explore mechanisms related to the "Developmental Origins of Health and Disease" (DOHaD) hypothesis. Some of these mechanisms appear to involve "nutritional programming" or genomic imprinting. Whether, and by what mechanism, early nutritional programming can be modified or reversed by exposures or interventions that occur later in the

[4] https://www.census.gov/newsroom/press-releases/2018/cb18-41-population-projections.html

child's life also requires further exploration. Because some epigenetic regulation may be paternal, it would be ideal if studies could obtain data from both parents.

Examples of future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Examine the impact of individual nutrient excesses or deficiencies and related contextual factors (e.g., weight status, inflammation, family stress, exercise behavior, water quality, ethnicity, health disparities) on measures of maternal health, fetal development, birth outcomes, and infant development.
- Generate new techniques to measure fetal nutritional status *in utero* more accurately.
- Perform mechanistic studies to determine how nutritional status of the parents before and during pregnancy affect the health and development of offspring and subsequent health and disease later in life.
- Explore the relationships between eating patterns and fertility in males and females.
- Identify genes and mechanisms responsible for the trans-placental fluxes of metabolites and nutrients and intergenerational transmission of health and risk for disease.

2-2. Research Nutrition in the Birth-to-24 Month Period

The first 24 months are integral to an infant's growth and development and can have lifelong health consequences. Compared to other parts of the lifespan, a greater number of knowledge gaps need to be addressed in the first 24 months. To systematically elaborate the relationship of nutrition with optimum growth and healthy development, information is needed about: 1) the dietary intake and nutritional status of the child during the 24-month period, 2) the composition of the child's diet, and 3) potential developmental milestones that can be used to describe optimal growth and healthy development.

2-2a. Assess the Influence of Diet and Nutritional Status on Infant Developmental and Health Outcomes

Studies are needed to clarify early life dietary exposures and nutritional status interrelationships with subsequent health and developmental outcomes. As mentioned in Priority 2-1, longitudinal studies are needed to examine these links. Such studies could provide new opportunities to refine the definition of optimal nutrition for the infant.

Knowledge about the mechanisms underlying food preference development among infants also is limited. How the order in which foods are introduced; the type, quantity, timing, and frequency of foods offered; and parental feeding styles affect food preferences, and whether and how these preferences persist, requires further investigation. Improved data on how and when complementary foods are introduced, including data on whether the infant was breastfed, formula-fed, or mixed-fed, will help researchers answer questions about early life exposures and the development of taste and food preferences that may play a role in subsequent dietary patterns. Additional knowledge gaps exist related to complementary feeding practices and the role of infant nutrition on the development of the human microbiome.

Future research activities that could be pursued in this Priority include but are not limited to the following:

- Examine the nutritional status of breastfed infants, and the age at which complementary foods are introduced, to determine the contribution of the complementary foods to any differences in nutritional status.
- Assess variability in potentially modifiable factors associated with nutrition-related health disparities, including maternal-infant feeding behaviors and beliefs about infant hunger and satiety to clarify these mechanisms and to inform interventions to prevent these disparities.
- Determine the impact of early exposure to salt and sugar on salt- and sugar-related food preferences and on dietary habits that are associated with chronic disease development in later life.
- Clarify the factors that lead to food allergies that develop in early life and develop efficacious dietary interventions to reduce the risk of food allergy.

2-2b. Enhance Knowledge of Human Milk Composition and the Biological Roles of its Components

Better data regarding the composition of human milk and the impact of different dietary patterns on milk composition would have many benefits. Dietary assessment and food composition data are typically used to determine nutrient exposures. Compositional data exist for many foods and infant formulas. However, many data gaps in human milk composition have recently been identified and should be addressed. Commercial infant formulas are modeled after human milk. Thus, studies that enhance knowledge about human milk composition and the relationship of its components to infant outcomes will likely benefit both breastfed and formula-fed infants.

Recent studies suggest that the concentrations or presence of some components in human milk may exhibit significant inter-individual variability. A better understanding is needed of the variability in individual components, what factors (e.g., including the impact of maternal diet) contribute to this variability, and what the implications of such variability are for the infant's immune system maturation, growth and development, and overall health. Obtaining a better understanding of these issues dovetails with efforts elsewhere in this *Strategic Plan* to examine inter-individual variability and personalized nutrition.

An increasingly common practice in the United States and other countries is for lactating women, particularly those working outside the home, to pump and store their milk for subsequent feedings or to donate to and/or use milk from "banks." Given the prevalence of this practice, more information is needed on whether and how storage, refrigeration, and freezing affects the nutritional composition, the physiologic function and potential health impacts of milk components.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Determine the physiological roles of breast milk components and the processes in the breast and other tissues involved in their formation and trafficking.
- Elucidate mechanisms through which human milk components influence an infant's growth and development and immune system maturation.
- Determine the role of human milk components in regulating tolerance to the gut microbiota.
- Quantify the extent to which milk components vary between individuals and within the same individual across the duration of breastfeeding. Identify the mechanisms responsible for variability.
- Determine the impact of different modes of storage on human milk components, including their nutritional value and safety.

2-3. Optimize Nutrition for Children and Adolescents

Only limited evidence is available on effective strategies and interventions to improve childhood and adolescent nutrition. Frequently, nutritional recommendations for this age group are derived from adult data. However, this may not be ideal as unique nutritional issues exist

during the periods of childhood and adolescence. Optimizing nutrition for children and adolescents is challenged by growth trajectories that may involve sudden growth spurts that change nutrient needs and can be compromised by both over- and under-nutrition. The period of childhood and adolescence is also an important time for brain development, and information about nutrition's impact on this critical developmental aspect cannot be elucidated simply by extrapolating data from adults.

Examples of future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Evaluate the role of foods and food patterns on optimal growth and development in childhood and adolescence.
- Explore the influence of specific foods or dietary patterns during childhood or adolescence on cognitive function and mental health.
- Conduct studies that test various dietary interventions during childhood and adolescence to examine their impact on health outcomes (e.g., cardiometabolic health).

2-4. Determine Nutritional Requirements for Healthy Aging in Older Adults

Nutrition is an essential component of healthy aging, although nutritional requirements for older adults are often not specifically investigated. Age-associated physiological changes may alter dietary intake and nutrient absorption. Thus, it is important to conduct research to understand how nutrient requirements change across the age spectrum and how these differing nutrient needs can be met. Inadequate micronutrient intake may be particularly important in the aging population from the perspective of metabolic control. In addition to dietary intake, physical activity and fitness have large influences on dietary intake and nutritional status as well as on health outcomes in older adults. Interactions between diet, nutritional status, and physical activity have been insufficiently examined.

Because a high percentage of older adults have one or more chronic diseases or ailments, research should examine the relationships and interactions between disease status, changing dietary intake, inadequate physical activity, chronic or acute inflammation, medication use, and advancing age. The simultaneous use of multiple medications and dietary supplements (polypharmacy) is more common among older adults than among other age groups and may

influence nutrient absorption. This issue may be particularly relevant during periods of acute illness or hospitalization, or in institutionalized individuals.

Another important issue in older adults is malnutrition, both under- and over-nutrition, as well as diet quality. Undernutrition in this population increases the risk of falls and cognitive impairment, and inadequate protein intake and sarcopenia are particularly significant problems. Research is needed to examine dietary intake, including consumption of specific food groups and dietary patterns, and the impact on age-related disability.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Identify specific nutrients and dietary patterns that reduce physical, physiological, and cognitive impairments.
- Identify unique factors and mechanisms (e.g., dentition, polypharmacy, mobility, nutrient absorption, thirst sensing) affecting nutritional status in aging.
- Conduct basic, interventional, or implementation research of nutritional approaches to reduce undernutrition in older susceptible populations, to offset age-related sarcopenia or to treat wasting associated with catabolic diseases and cachexia.
- Determine the impact of inadequate dietary intake in older adults on frailty and morbidity during hospitalization.

2-5. Determine Mechanisms by which Dietary Patterns Affect Health Status and Chronic Disease Susceptibility

This Priority focuses on increasing knowledge about how specific patterns of eating influence health outcomes and chronic disease susceptibility. Methodological opportunities in the dietary pattern field related to this topic are described in Priority 6-3. A great deal of research has been dedicated to studying the physiological effects of individual nutrients. In recent years, the focus of research has expanded to examine the impact of total diet on health and disease outcomes. Dietary patterns provide a comprehensive way to characterize dietary exposures and explore diet and health relationships. This Priority explores research needs as they relate to dietary patterns characterized by food groups consumed, timing of when food is consumed or frequency of food consumption, and amount of food ingested. Patterns involving timing and frequency often involve periods of caloric restriction or fasting. Further research is

needed to examine how dietary patterns produce beneficial or detrimental effects on health status and how these effects may change across the lifespan. Dietary patterns are multidimensional, representing both a multi-layered exposure and a behavior, as well as dynamic in that patterns vary over time and across the life stage.

2-5a. Provide Mechanistic Insights into Dietary Patterns and Chronicity to Reduce Chronic Diseases

Foods and nutrients are consumed in a variety of combinations and can have interactive and potentially cumulative effects on health status. Dietary pattern research may focus on specific types of diets (e.g., Western, Mediterranean, vegetarian). Not only do these patterns have either deleterious or beneficial effects on health, depending on individual response, they provide a comprehensive way to capture the interaction of nutrients and bioactive compounds within the whole diet. Mechanistic and longitudinal studies are needed to explore the mechanisms by which dietary patterns affect physiology and metabolism and the impact that different dietary patterns have on health promotion and disease prevention. These insights can provide information from which to derive future dietary guidance.

Research designed to study the impact of dietary behaviors that involve the frequency of food consumed or the timing of when food is eaten may also provide valuable insights. Energy metabolism is not static and may require precise coordination of behavior, physiology, and molecular process across the 24-hour cycle. Mounting evidence links circadian misalignment to an array of adverse health conditions, including obesity, diabetes, hypertension, stroke, cancer, and GI disorders. Thus, research suggests that behaviors such as the number of meals or eating occasions per day, meal skipping, and hours between eating occasions and/or fasting all have biological consequences. Further investigation is needed to elucidate the pathways behind these effects.

Caloric restriction may also have health benefits for some adults. Prolonged caloric restriction without malnutrition has been linked to extended lifespan and delays in age-related diseases. Several mechanistic studies have identified multiple signaling pathways known to modulate the aging process that are affected by caloric restriction (e.g., oxidative stress response, damage repair, inflammation, autophagy, proteostasis, and nutrient sensing). However, further studies are needed to fully explain these associations.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Intensively and systematically characterize the effects of specific dietary patterns on physiological, behavioral, and biological measures or biomarkers, and determine the mechanism(s) responsible for emerging differences.
- Explore methods to define dietary pattern trajectories across the lifespan and identify transition points that are most amenable to intervention.
- Determine the efficacy of time-restricted dietary patterns, with or without calorie restriction, to prevent and/or manage metabolic-related chronic diseases or to promote healthy aging.
- Test the long-term effects of interventions leveraging biological rhythms (e.g., circadian rhythm) and/or caloric restriction on diet-related disease risk or surrogate markers of aging in clinical trials or natural experiments.
- Determine how meal habits (e.g., three meals per day versus frequent snacking) developed in childhood alter life-long patterns of food intake and the ability to maintain a healthy body weight later in life.

2-5b. Leverage Clinical Trials, Natural Experiments, or Other Rigorous Study Designs to Examine How Dietary Patterns Affect Health Outcomes

As illustrated in Priority 2.5a, identifying the underlying mechanisms by which dietary patterns confer health benefits or deleterious outcomes could provide expanded opportunity for targeted interventions. In addition to further mechanistic research and prospective and high-quality randomized controlled trials, another way to explore the causal associations of dietary patterns with health outcomes is to use alternative designs, such as natural experiments, especially in cases where controlled experimentation would be prohibitively expensive or difficult to implement.

Alternative or quasi-experimental designs have been increasingly recognized for their ability to enhance research in the real world. Natural experiments are empirical studies in which individuals are exposed to the experimental and control conditions as determined by nature or by other factors outside the control of the researchers but where the process governing the exposures resembles random assignment. Natural experiments are most useful when investigators are examining a clearly defined exposure in a well-defined subpopulation (and the

absence of an exposure in a similar subpopulation) such that changes in outcomes may be plausibly attributed to the exposure. To make progress in the use and success of natural experiments, methodological and analytic advances are needed to strengthen the evaluation of these studies. Taking advantage of natural experiments and other alternative research designs could help to elucidate the protective benefits of various dietary patterns.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Conduct prospective and high-quality randomized controlled trials, including pragmatic trials, aimed at determining how eating patterns influence health and life span outcomes.
- Perform ancillary studies to examine mechanisms by which specific dietary patterns are thought to elicit health benefits.
- Elucidate biomarkers of adherence to specific dietary patterns.
- Perform comparative effectiveness research using patient registries or other large-scale data resources to evaluate impact.
- Examine the role of social factors, including those associated with health disparities, on dietary pattern selection, energy balance, health status, or diet-related disease prevalence.
- Consider how dietary patterns that have demonstrated positive health benefits can be adapted for different cultures, socioeconomic groups, geographic locations, and health conditions.

Theme 3. Explore Individual Variability in Response to Diet Interventions to Inform Nutrition Science, Improve Health, and Prevent Disease

Introduction

Interest in developing “precision medicine” interventions to prevent, treat, and manage disease is growing. The idea of precision medicine has arisen from observations that some treatments may work better in some individuals than in others. Although such variability in therapeutic responses was at first discouraging, these situations are now thought of as scientific opportunities to identify and recognize previously unaccounted for or unrecognized factors affecting efficacy.

Such inter-individual variability in responses extends to nutrition interventions, even in genetically identical preclinical models. A wide range of factors is thought to have a role in an individual’s physiologic response to diet and food components. These include, but are not limited to, genetics, epigenetics, metabolism, gut microbiota diversity and metabolism, inflammatory status, sleep and exercise habits, environment (e.g., food access and exposure to toxins and/or pathogens), social and health disparities, socioeconomic factors, medication use, and the presence of disease or conditions. A tremendous opportunity exists to improve knowledge about how distinct dietary patterns and intake interact with these factors, and to study the comparative role and interaction of these factors on individual health outcome responses to dietary interventions. A key priority in this area is to test the efficacy and effectiveness of personalized intervention approaches tailored to individuals' genetic variants, microbiome, metabolic profiles, and above-mentioned or unknown factors compared with traditional "one size fits all" approaches in improving dietary quality and health outcomes. Overall, researchers need a better understanding of when population versus targeted/precision nutrition interventions will be optimal.

Sophisticated bioinformatics and computational approaches will be needed to evaluate the interaction, additivity, or synergism of potentially responsible factors and systems. Doing so will often require use of multiple, large datasets often referred to as “Big Data.” However, the concept of Big Data is more than just very large data or multiple data sources and encompasses the complexity, challenges, and new opportunities presented by the combined analysis of data. Big Data studies can rely on information from massive datasets or data that are generated continuously and with great diversity or uncertain quality. In biomedical research, Big Data

sources can include the diverse, multimodal data being generated by researchers, hospitals, and mobile devices around the world.

Applying innovative biomedical Big Data approaches and advanced analytics in human nutrition research has great potential to accelerate understanding of individual variability in response to diet. However, for this potential to be realized, integrating data from multiple sources is required. For example, to explore factors influencing individual intake, biomedical imaging, genetic, phenotypic, and molecular data could be integrated with data on the food supply, food purchasing and acquisition, and on household food security. Models of individual consumption behavior require data about individual behavior as well as factors influencing such behaviors. Developing system science approaches, methods, tools, and the interfaces that make Big Data resources more useful and accessible to the nutrition research community are needed to facilitate these avenues of research.

Scientific Priorities

3-1. Elucidate the Biological Factors Underlying Individual Variation in Response to Dietary Interventions

Biological factors contribute to inter-individual variability in dietary choices (e.g., when, how much, and what is eaten), nutritional status, and the physiological and health consequences of diet. Major gaps exist in current knowledge about these factors, how they interact, and exactly how they contribute to variability in diet-health relationships.

Relevant biological factors include, but are not limited to, genetics, epigenetics, gut microbiota metagenomics, metabolomics, transcriptomics, anthropometrics, behavioral phenotype, chronotype, heart rate, GI transit times, and the presence of inflammation. These factors can be assessed with various tools, including health histories, clinical laboratory measures, electroencephalography, bioimaging, and sensor data. Whereas some behaviors affected by biology occur at the level of consciousness, others, such as hedonic responses to taste, salivation, rates of gastric emptying, incretin release and responses, and digestive responses occur subconsciously. Understanding the role of these factors in inter-individual responses to diet and nutrition is important. Of equal importance is the need to improve the current standards and tools used to assess nutritional status to determine normal intra-individual and inter-individual response to and metabolic utilization of some nutrients and dietary components. Establishment of such metrics would allow comparisons across diverse populations.

Research relevant for this Priority include mechanistic studies to explore how the above-mentioned biological factors influence inter-individual variability in cases where that influence is not obvious. Experimental evidence or computer learning algorithms or modeling could be deployed to provide these mechanistic insights.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Identify potential biomarkers or biomarker patterns that could help to stratify heterogeneous populations relative to their dietary responses.
- Determine how diet, genes, and other factors affecting nutrition inter-relate and can be used to predict dietary responses or health outcomes and determine whether these can be leveraged to improve health.
- Develop and test mechanistic hypotheses explaining how biological factors lead to inter-individual variability.
- Examine targeted intervention strategies (biologic, behavioral, pharmacologic, and environmental) to help guide optimal dietary approaches.
- Explore use of alternative statistical approaches that examine individual variability in response to interventions in addition to use of more traditional approaches that focus on estimation of group means.

3-2. Examine the Contribution of Social Determinants of Health and Other Environmental Factors to Variations in Response to Nutrition Interventions

Social determinants of health include age and the conditions (e.g., social, economic, and physical) in the environments in which people are born, live, learn, work, play, worship, and age that affect a wide range of health, functioning, and quality-of-life outcomes and risks, including nutritional status. Two such conditions include access to healthy foods and food insecurity. Food insecurity has direct and indirect consequences, including poor dietary intake, poor physical and mental health, hospitalizations, stress, reduced academic achievement, and fetal epigenetic changes. These consequences alter nutrition and health relationships in ways that need to be better elucidated and integrated with other interacting factors in this Theme. In addition, certain populations, including African Americans, Hispanics, and Native Americans,

exhibit a higher incidence and prevalence of, and mortality from, diet-related chronic diseases. Given the disproportionate burden of disease and disease risk factors, and frequently disparate access to nutrient-dense food and health care, it is vital to understand how nutrition differentially affects health risks and outcomes in these groups. Understanding the dietary preferences and practices of an increasingly diverse and aging population is necessary for clinical and community-based organizations if they are to provide personalized, tailored, and equitable dietary recommendations and resources.

It is important for researchers, clinicians, and other health stakeholders to understand how social determinants of health and other environmental factors contribute individually and in combination to inter-individual variability in the relationships between diet and health. Multi-level interventional and systems research could help to develop guidelines for improving diet and other health-related behaviors.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Elucidate the role and significance of biological, psychosocial, environmental, and sociocultural factors in inter-individual or inter-population variability in responses to diet interventions, dietary patterns, or nutritional challenges, along with their implications for health and disease susceptibility.
- Determine the mechanisms underlying the co-existence of food insecurity, obesity, and other related metabolic conditions.
- Investigate how to leverage affordable mobile technologies and applications (“apps”) to change consumer purchase behaviors and improve diet quality among diverse individuals
- Elucidate how diversity of all kinds (e.g., race, ethnicity, acculturation, age) affects nutrition and health interrelationships.
- Apply machine-learning algorithms and simulation models incorporating biological, psychosocial, environmental, and sociocultural factors to identify potential responders and non-responders to diet interventions to predict improvements in diet-health interrelationships.

Theme 4: Enhance Clinical Nutrition Research to Improve Health Outcomes in Patients

Introduction

The focus of this Theme is to advance knowledge of how nutritional support and interventions for patients can be leveraged to improve the treatment, management, and recovery from specific diseases. Research provides evidence of an association between poor nutritional status and prolonged length of stay in hospital, decreased quality of life, and increased morbidity and mortality of patients in clinical settings. Correcting nutrient deficiencies and malnutrition in hospital and outpatient settings is expected to help decrease comorbidities, improve recovery times, and reduce readmission rates. Two key challenges for clinical nutrition include: (1) limitations in current objective biomarkers for malnutrition (including macro and micronutrients), and (2) the relatively modest extent of evidence-based research on diet and nutritional support interventions and relevant treatment endpoints for patients with acute or chronic health conditions.

A critical aspect of the first key challenge is the inherent difficulty in assessing nutritional status and screening for malnutrition. Such screening can be complicated by the presence of underlying chronic or acute conditions, such as inflammation. These conditions often synergistically interact with and affect traditional measures of nutritional status. Thus, overcoming the first challenge in performing research is to identify better biomarkers of nutrition to assist in the appropriate identification of specific patient subgroups that may benefit from nutritional interventions.

The second key challenge is related to the many cases in which inpatients or outpatients require nutrition support therapy, including oral, enteral. or parenteral nutrition. Many of the studies on which professional association guidelines are based have limitations in terms of sample size, patient heterogeneity, variability in disease severity, lack of baseline nutritional status, or insufficient statistical power for analyses. Research is needed to elucidate what types of nutritional support interventions are needed for which patients and during which periods of active treatment and recovery and to clarify which clinical endpoints can be used to determine treatment success. The identification of specific needs for specific subpopulations of patients is likely to require research within the context of large collaborative research initiatives in large clinical practice systems and institutions. Research is needed to accurately quantify use, outcomes, and impact of current nutritional interventions. Such data, when analyzed by high

dimensional biostatistical and computational programs, might be very helpful for hypothesis generation to develop new quality-of-life and health improvement initiatives.

This Theme also addresses specialized clinical nutrition support and services that are needed for individuals with unique nutritional needs. These individuals include premature infants, individuals with inborn errors of metabolism, and individuals with disease-associated alterations in nutrient absorption or metabolism. In many cases, there is only a limited evidence base to inform the delivery of clinical nutrition services or the development of specific medical foods for these individuals.

Scientific Priorities

4-1. Identify and Leverage Interactions Among Nutrition, Disease States, and Treatments

Understanding how various disease states or conditions affect dietary intake and how these disorders can be addressed is important to improve patient health and recovery. Many ailments affecting the oral or digestive tract can impair dietary intake. Examples include missing teeth, orofacial pain conditions, salivary dysfunction, oral complications during cancer treatment and certain GI disorders. Other conditions can decrease desire to eat or the sensation of hunger, such as cancer cachexia, GI diseases, drug side effects, and infections. Lastly, other diseases may lead to malabsorption or altered dietary patterns leading to nutrient deficiencies. Recognition of how disease conditions change dietary patterns, how these changes affect treatment, and the identification of effective interventions to address potential nutritional deficiencies would have a positive impact on patient health outcomes.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Elucidate the mechanisms by which specific diseases or conditions alter dietary intake or nutritional status and determine how nutritional status influences the effectiveness of treatment on the progression or treatment of disease.
- Determine optimal dietary intakes of nutrients and foods to improve glycemic control and reduce cardiovascular risk factors among patients with metabolic diseases.

- Identify nutritional approaches to improve cancer survival, especially among individuals with breast cancer, colorectal cancer, or prostate cancer.

4-2. Improve Assessment of Malnutrition

This Priority seeks to develop improved approaches for assessing nutritional status, especially for use in clinical situations where malnutrition can interfere with optimal treatment of, or recovery from, disease. Malnutrition is usually thought of as affecting populations in countries with limited resources and typically characterized by not having enough to eat (undernutrition). However, individuals in the U.S. are also at risk, and the term malnutrition refers more comprehensively to deficiencies, excesses, or imbalances in a person's intake of energy and/or nutrients. Approaches for characterizing deficiencies in specific nutrients (or their imbalances) and energy need to be refined. For energy excess, body mass index (BMI) is used as a screening tool (albeit overweight and obesity may be associated with micronutrient deficiencies). Further research is needed to improve or replace BMI, as it can be misleading in specific life stages and certain body types. Overall, additional research is needed to improve diagnostic tools and protocols for screening and diagnosing the various forms of malnutrition that can negatively affect medical and surgical outcomes.

A variety of risk factors are associated with undernutrition. Socioeconomic and environmental factors, such as food insecurity and limited food access, can increase an individual's risk of becoming malnourished. Acute and chronic diseases (e.g., cancer, stroke, chronic obstructive pulmonary disease, heart failure, infection), inflammation, and conditions in which individuals cannot feed themselves due to trauma or surgical procedures also are risk factors.

An important goal of this Priority is to support the establishment of, and provide supporting evidence for, objective and cost-effective biomarkers of energy, protein, and micronutrient-related undernutrition or imbalances that are not affected by or can be adjusted to account for the presence of other confounding conditions, such as inflammation.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Identify objective, standardized methods or markers for assessing and diagnosing malnutrition that would be broadly useful for clinical care and/or can be adapted for specific diseases states or personal characteristics.
- Facilitate the translation of research-derived biomarkers to clinical practice.

4-3. Identify Triggers and Endpoints for Nutritional Support in Clinical Settings

Patients requiring medical care may present with or develop undernutrition and/or other barriers to sufficient oral intake. In some cases, patients may require enteral or parenteral nutrition support or supplements to assure appropriate and adequate nutrient intake depending on their health status. The evidence base to establish when these treatments should be implemented, how they should be managed, and when they should be discontinued is limited. Standardized and/or core measures to evaluate clinical nutrition support implementation and clinical outcomes are also needed. Furthermore, health status, indication for nutrition support, and personal characteristics (e.g., age, BMI, nutritional status) are important factors to consider.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Identify appropriate objective and evidence-based measures or endpoints that can be used to guide the initiation and cessation of clinical nutrition interventions to improve quality of care, health, and recovery in specific disease states and during pre- and post-operative care.
- Determine the optimal approach to monitoring tolerance and adequacy of nutritional support therapies.
- Examine tissue or organ dysfunction in response to nutritional therapies.
- Elucidate the impact of nutritional support interventions on disease morbidity, progression, recovery, surgical morbidity and recovery, and health care costs.

4-4. Optimize Nutritional Support for Premature and Low Birthweight Infants

About 10 percent of the infants in the United States are born prematurely.[5] Premature infants have underdeveloped GI tracts and may initially require parenteral, enteral, or other nutritional support. Research gaps remain regarding the appropriate approaches to feeding premature infants and caring for their associated health problems. For example, the optimal ratios of macro- and micronutrients for premature infants are not known. In many cases, the mother's or a donor's breast milk is used, but that may not sufficiently meet the needs of these vulnerable infants.

Research to optimize nutritional support for premature infants has three main goals. The first is to better understand their nutritional requirements and the development of their GI tracts. The second goal is to develop a stronger evidence base to inform nutritional guidance for this patient group. The third goal is to develop and codify tailored nutritional strategies for premature infants, which may open the door for precision nutrition for infants across a spectrum of birth gestational ages and postnatal stages throughout their hospitalization and afterward during critical stages of development.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Determine how nutrient digestion and absorption is affected by adjusted age and coexisting morbidities of prematurity.
- Develop the evidence base to inform nutritional practices that are linked to specific markers of intestinal development.
- Identify the optimal macronutrient and micronutrient intakes associated with long-term outcomes (e.g., improved growth, bone mineralization, neurodevelopment) in preterm infants of different adjusted ages.
- Determine whether commonly used clinical measures (e.g., anthropometric measurements, hematology, and clinical chemistry values) are sufficiently sensitive and specific to assess

[5] https://www.cdc.gov/nchs/pressroom/sosmap/preterm_births/preterm.htm

nutritional status (i.e., sufficient, marginal, or deficient) during hospitalization and after discharge.

- Establish best practices for determining enteral feeding readiness, to assist in transitioning the preterm infant to breastfeeding and overcoming barriers to successful breastfeeding.
- Examine the efficacy of and developmental effects in preterm infants of donor milk, milk fortification, trophic feedings (providing small amounts of enteral food to stimulate development of the GI tract), routes of feeding, and feeding timing or type (e.g., specific periods of the day, bolus versus continuous).
- Develop best practices for feeding preterm infants with common infant disorders (e.g., patent ductus arteriosus, acute illness, malformations of or damage to the GI tract from congenital defects, biliary atresia, necrotizing enterocolitis, or sepsis) or who receive therapeutic interventions (e.g., transfusions).

4-5. Optimize Nutrition Interventions for Patients with Special Nutritional Requirements

Individuals with inherited diseases and diseases associated with loss of organ function or altered handling of nutrients may have distinct nutritional requirements. These diseases and conditions include inborn errors of metabolism, food allergies, gastrointestinal disorders, and chronic kidney disease. Additionally, certain diseases may require special monitoring of dietary components (e.g., sodium in heart failure). Although some patients with these diseases can maintain their nutritional health through counseling and by modifying their normal diet, other patients require other types of nutritional therapy or specialized medical foods to maintain their nutritional health. Additional research is needed to build evidence-based nutrition support recommendations for individuals with these disorders.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Develop specific biomarkers for morbidities associated with inborn errors of metabolism and other diseases involving special nutritional considerations.
- Identify best practices and overcome barriers for providing nutrition support for special populations, including those with inborn genetic errors or loss-of-function diseases.

Theme 5: Advance Implementation Science to Increase the Use of Effective Nutrition Interventions

Introduction

It is important to support the development of evidence-based nutrition strategies to improve health. Recognized and significant gaps exist between discovery and broad adoption of proven-effective nutrition strategies. Research to identify barriers to adoption and to test proven solutions to overcome them is required. Another area of opportunity is to equip health care providers, patients, families, caregivers, and communities with tools to adapt and sustain effective nutrition practices. Implementation science, the study of methods to promote the adoption and integration of evidence-based practices, interventions, and policies into routine health care and public health settings to improve the impact on population health, is therefore a priority described in this *Strategic Plan*.

Evidence-based nutrition strategies should be implemented across the lifespan to change dietary behaviors. Many healthy dietary behaviors have been identified, and dietary guidelines have been published recommending these behaviors. However, few of these behaviors have been broadly adopted by the public. For example, over the past 20 years, numerous clinical trials have demonstrated the health benefits of reduced sodium/high potassium diets, but widespread adoption of these recommendations by the public has been limited. Research to discover effective approaches to enhance widespread adoption of these recommendations is needed.

Many frameworks (e.g., social and ecological) suggest dietary behaviors are shaped at multiple levels and through interactions across these levels. The intersection of biological, psychosocial, sociocultural, and environmental factors may act additively or synergistically to improve or decrease the chances that individuals will change their dietary behaviors and sustain those changes over time. Thus, intervention strategies targeting multiple levels of the food environment while integrating behavioral phenotypes and biological factors are needed. The potential to reveal, or track, implementation science strategies, suggests that investments in Big Data, systems science research, and other meta-analytic strategies that can assess the contribution of multilevel intervention strategies may help to sustain behavioral change. Additionally, thoroughly phenotyping (biologically and behaviorally) participants in the context of their environmental exposures may provide insights that could be used to better refine interventions for responders and to design new approaches for non-adopters.

Understanding the barriers to changing dietary behaviors, sustaining adherence to those changes over time, and developing scalable interventions will require advances in the basic science of behavior change. Increasing knowledge about behavioral phenotypes and the relative importance of competing or facilitating biological and multi-level environmental factors also may help target implementation science strategies. Equally important is the use of different implementation science designs and methodologies beyond randomized controlled trials of efficacy. Designs that are readily adaptive (e.g., hybrid designs) and include a strong evaluation component are essential for effective interventions in nutrition research.

Scientific Priorities

5-1. Evaluate Strategies for Sustaining Nutrition-Related Behavior Change

The *Dietary Guidelines for Americans* provides evidence-based recommendations for a healthy diet.[6] However, only limited evidence is available on how to change intake behavior to achieve healthy dietary patterns and sustain those patterns over time. Long-term studies in diverse populations are needed, involving interdisciplinary research teams of psychologists, registered dietitians, physicians, and other healthcare professionals, communication strategists, and individuals from other fields and focusing on understanding human food purchasing and eating behavior.

Most research in this area has been conducted in simulated food environment settings and may not accurately reflect real-world behavior. Thus, research is needed to determine which combination of strategies and conditions has the greatest potential to foster healthy eating in real-world settings. Research at the individual, interpersonal, community, and societal levels is needed to inform efforts to achieve sustainable dietary behavioral change.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

---

[6] U.S. Department of Health and Human Services and U.S. Department of Agriculture. *2015 – 2020 Dietary Guidelines for Americans*. 8th Edition. December 2015. Available at https://health.gov/dietaryguidelines/2015/guidelines/

- Identify the personal and social factors that mediate eating behavior and that may be partially responsible for diet-related health disparities (e.g., heart disease and obesity).
- Determine the multilevel influences on dietary patterns through research examining the food supply, industry practices, advertising, and the retail food environment.
- Consider the interplay of multiple health-related behaviors that tend to cluster together (e.g., tobacco use and diet, and physical activity and diet) and examine whether changing the non-dietary behavior influences nutrition-related behaviors.
- Explore of the impact of nutrition and cooking literacy on adherence to healthy dietary behaviors or diet recommendations for specific forms of chronic disease.
- Explore how to modify organoleptic properties of foods to improve adoption of effective dietary interventions.
- Leverage natural experiments to evaluate the effects of changes in the food system, including issues such as labeling, prices, and availability.
- Examine implementation of dietary guideline recommendations across the lifespan.

5-2. Leverage Multidimensional Strategies to Increase Use of Evidence-Based Nutrition Approaches

The challenges of attaining widespread adoption of evidence-based dietary interventions have heightened the recognition of the need for implementation science research to: (1) identify how to effectively promote the adoption of healthy dietary patterns in specific population groups, (2) test, evaluate, and implement cost-effective and efficacious approaches to implement and deliver nutrition or dietary interventions, and; (3) enhance the sustained use or adherence to healthy dietary choices.

Research strategies are needed for effective and sustained delivery of nutrition interventions within the context of health care, including collaborating with many different types of healthcare providers and enhancing interoperability of data collection systems (e.g., electronic health records) in healthcare and other community-based implementation settings. Overall, implementation and effectiveness research will help identify approaches for broadening the adoption and impact of population-level efficacious nutritional interventions.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Use systems science and other methods [e.g., sequential multiple assignment randomized trials (SMART) and hybrid designs], to estimate how purposeful changes designed to improve population health within local, state, tribal, or federal food and nutrition policies and programs could improve diet, health, and social outcomes.
- Investigate the role of health professionals (e.g., registered dietitians, lactation consultants, nurse practitioners) in collecting data and conducting outreach or interventions.
- Explore methods, including community engagement, to develop and disseminate nutrition messages to target populations and specific communities within these populations (e.g., based on age, race/ethnicity, sex, geography).
- Advance behavior change theory by examining how biological, psychosocial, environmental, and sociocultural factors affect the differences among individuals in their nutrition behaviors in response to nutrition interventions.
- Examine the effectiveness of large-scale policies and programs that could help improve diet quality at the population level or identify any differential effects on sub-populations (e.g., financial incentives and disincentives, such as taxes on purchasing behaviors).

Theme 6: Develop and Refine Research Methods and Tools

Introduction

Theme 6 provides a vision for transformational innovations in nutrition research methods. These include development of diet and nutrition assessment methods, identification of biomarkers of nutrition and diet-related chronic disease or surrogates, and ways to better understand and address the systems that affect and are affected by nutrition. Regarding biomarkers of nutrition, these include sensitive and specific measures of nutrient exposure, status, or function.

This Theme encourages new and ongoing efforts to improve and reduce measurement error in existing and emerging tools and methods. Improving existing methods and developing new statistical approaches to enhance rigor, as well as computational approaches to make them more accessible, are also priorities. When tools are developed or available, strategies may be needed to ensure their usability and adoption by the research community. Overall, these cross-cutting methods and tools will benefit the research community and enhance research prioritized in previous Themes.

Scientific Priorities

6-1. Improve Methods to Assess Dietary Intake and Exposure

Advances have been made in recent decades in the field of dietary assessment, which have facilitated research to identify links between diet and health. Nevertheless, a consensus exists in the research community that this is an area where nutrition research should be strengthened. Accelerating “efforts to improve nutrient and other food substances exposure assessments” was one of the primary recommendations of a 2017 National Academies of Sciences, Engineering, and Medicine report on *The Development of Guiding Principles for the Inclusion of Chronic Disease Endpoints in Future Dietary Reference Intakes.*[7]

[7] National Academies of Sciences, Engineering, and Medicine. 2017. *Guiding principles for developing Dietary Reference Intakes based on chronic disease*. Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/24828.

The current state-of-the-art approach to capture dietary intake relies primarily on self-report. Ongoing efforts to improve self-report methods to reduce measurement error are important priorities. In addition, it would be beneficial to develop new objective methods or technologies for collecting dietary intake data and to discover new biomarkers that have potential to both capture and validate self-reported intake data. Although it seems unlikely that it will be possible to completely replace self-report methods, approaches that leverage new methods and technologies could help improve dietary assessment. To the extent possible, self-reported dietary intakes should be validated against objective biomarkers versus other error-prone self-report dietary intake instruments. Identifying new methods that use diet-dependent biomarkers and other indicators has the potential to transform the field of nutrition research as it relates to biological response.

Enhanced statistical methods (e.g., to reduce measurement error) and computational tools will be needed to analyze the complex datasets that might emerge from new and improved dietary assessment methods. These will also be needed to appropriately merge self-reported data with other data from new technologies.

6-1a. Encourage the Development and Use of New Technologies to Capture Dietary Data

Capturing information about the totality of diet—including the timing of meals and snacks, portion sizes, combinations of foods, and preparation methods—is a central need of nutrition research. It is important to develop or leverage emerging technologies to augment self-report methods to address issues of bias and stigma which are known to effect self-reports.

Innovative technology-based devices and methods will need to minimize measurement error, cost, respondent burden, and subjectivity while at the same time improving precision and accuracy. One of the goals and opportunities in this area is to develop fully scalable tools, involving systems approaches and methods that can be applied to millions of people at low cost, possibly helping to transform nutrition research by allowing for more precise dietary intake assessments.

Another aspect to consider in developing new methods for dietary data capture is the level at which the data are collected. Much of dietary assessment and diet-related research focuses on the individual. However, it is critical to capture information about the broader food

environment, as well as about an individual's food availability and consumption in the home, school, or workplace.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Use new technologies to develop innovative and culturally appropriate approaches for capturing dietary intake data on the individual, family, household, community, and societal levels.
- Discover or refine technologies involving active and passive data capture using "smart" hand-held devices, wearable devices, or other technology that can automatically collect data on eating habits.
- Leverage mobile, home, or other devices to augment recall data or to integrate with other experimental measurement technologies.

6-1b. Develop Biomarkers to Determine Exposure to, and Quantitative Estimates of, Usual and Recent Intake of Foods, Nutrients, and Dietary Constituents

Another approach for assessing dietary intake makes use of laboratory biomarkers. There is a need to validate dietary intake assessment methods that correlate strongly with objective biomarkers, not different questionnaire- and recall-based methods. The most important role of such biomarkers is to adjust for measurement error in self-report and other dietary assessment instruments prone to biases. Encouraging the development and evaluation of biomarkers of dietary intake to augment self-reported measures of recent and usual dietary intake is a key aspect of this Theme. Ideally, biomarkers should represent objective, accurate, and unbiased measures of relatively stable relationships with dietary intake that can feasibly be used in large studies.

For most nutrients or dietary constituents, an exposure biomarker has not been discovered or characterized. Further research is needed for biomarker development spanning discovery, early characterization, initial validation, generalizability, and refinement. The most important role of dietary biomarkers is their use to adjust for measurement error in self-reports and other instruments prone to biases in dietary assessment.

It is unlikely that new dietary assessment methods, even if objective, would be free of biases and it is unlikely that biomarkers will be available for all foods, nutrients, and dietary patterns of

interest. The promising role of biomarkers is, therefore, to evaluate and adjust for measurement error in assessing the corresponding dietary components and to also assist in conducting sensitivity analyses for the effect of measurement error in assessing other dietary components. The existing biomarkers would provide a range of different measurement error components (e.g., intake related bias, person-specific bias, and random within-person error) to conduct an appropriate sensitivity analysis with realistic assumptions.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Conduct research aimed at discovery of new, unbiased biomarkers of intake.
- Model the kinetics (i.e., time course after a single dose) of existing or newly discovered biomarkers and dose-response characteristics.
- Assess how individual characteristics of research participants (e.g., body weight, chronic disease, age, usual intake of a nutrient(s), and food sources or preparation methods) affect biomarker abundance and kinetics.
- Validate biomarkers in subpopulations or in the context of certain health conditions (e.g., urinary potassium excretion among African Americans, urinary sodium excretion in people with kidney disease.)
- Conduct studies to advance the use and scalability of biomarkers for diverse populations and environments.
- Use “omic” (e.g., metabolomics, genomics, transcriptomics, proteomics) approaches in short-term intervention studies and long-term observational studies to identify and validate potential biomarkers of a dietary pattern, food group, or food component of interest.

6-1c. Improve Statistical Methods and Computational Solutions for Dietary Assessment

Assessing diet with accuracy and precision is particularly difficult and presents many opportunities for analytical variation and measurement error. The field of dietary assessment research has acknowledged this issue, identified the major types of measurement error, and made important strides in addressing them. Measurement error may be systematic, random, intentional, or unintentional (e.g., poor nutrition literacy, prevailing norms). Accounting for

measurement error is important in the context of both longitudinal epidemiological or observational clinical studies that capture the multivariate, dynamic nature of diet and interventional clinical trials where identifying change in consumption over time and between groups is critical. Thus, a priority of this *Strategic Plan* to address measurement error both in existing methods of dietary assessment and new methods that may emerge.

Modeling and corresponding computational approaches will be important to address dietary assessment needs for populations and comparison of selected groups of individuals. Accounting for individual variation through computational approaches is needed to examine relationships between diet, nutritional status, health, and the various factors that affect those relationships. Computational methods and machine learning can help to evaluate biomarkers for phenotypes of dietary classifications, intake of nutrients, foods, and other dietary components.

New methodologies will also be needed to help investigators analyze complex datasets, use existing and emerging statistical methods to reduce measurement error, and merge different types of data for dietary assessment. As these methods and computational tools are developed, it will be important to ensure that they are usable and widely available to the broader research community.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Model measurement error in potential dietary intake biomarkers to improve their predictive value.
- Develop robust measurement error correction procedures that can be widely used in large population-based studies.
- Create new technologies or refine existing approaches that address validity and sensitivity.

6-2 Develop and Validate Biomarkers of Relevance for Chronic Disease and Diet

In many cases, long periods of time precede the appearance of a diet-related chronic disease. This makes it difficult to determine a link between nutrition-related exposures and disease. There is a need to identify and validate surrogate biomarkers that can be used to

predict chronic disease risk, similar to the use of elevated blood pressure or proteinuria as surrogate biomarkers for future cardiovascular disease or kidney disease, respectively. Chronic disease biomarkers should reflect a biological process, and if used for clinical purposes, they should satisfy criteria recommended by the National Academies of Science, Engineering, and Medicine.[8] Such biomarkers are not limited to plasma clinical chemistries, but could involve histological, cell, or tissue-based imaging or physiological measures. Potential candidate biomarkers may be at different phases of the development spectrum, with some needing discovery research support, and others requiring further characterization, testing, or refinement.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Investigate the validity, reliability, and reproducibility of new potential or existing chronic disease biomarkers of nutritional status or intake.
- Validate biomarker performance across different populations.

6-3. Improve Methods for Dietary Pattern Analysis

Nutrition research has been shifting away from the reductionist approach of investigating individual foods or nutrients and disease toward a more comprehensive approach to studying the health effects of food combinations and the timing, amount, and frequency of food consumption. Thus, many studies attempt to identify eating behaviors and dietary patterns that may contribute to the increased or reduced risk of diet-related diseases. The methods to investigate each of these different types of dietary patterns need improvement. For example, dietary pattern analysis is hampered by the lack of tools that can capture the multi-dimensional, dynamic, and longitudinal nature of diet along with factors that might influence dietary intake, such as psychosocial, sociocultural, and environmental factors. Another challenge is comparing the results of dietary pattern analyses across studies because patterns are not consistently defined or labeled. Developing consistent approaches and practices would facilitate meta-

[8] National Academies of Sciences, Engineering, and Medicine. 2017. Guiding principles for developing Dietary Reference Intakes based on chronic disease. Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/24828.

analyses and cross-study comparisons. Increased standardization of the constructs and methods used to define dietary patterns will help inform studies that examine the impact of these dietary patterns on biology and various health parameters across different life stages.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Improve methods to precisely characterize dietary patterns.
- Develop and validate markers or signatures of certain dietary patterns across the lifespan in healthy individuals and those with specific diseases.
- Develop tools and statistical approaches for incorporating other aspects of dietary patterns, such as timing and frequency of food consumption, into dietary pattern analysis.
- Develop tools that can be used to assess food patterns beyond individuals or nutritional quality, for example in supermarkets, schools, restaurants, and households.
- Develop a comprehensive approach to capture multiple factors influencing individual diets and dietary patterns.
- Develop tools to assess the multi-dimensional and dynamic nature of diet.
- Extend efforts to standardize methods and indices used to define dietary patterns and establish a conceptual framework for eating patterns to improve the translation of findings for policy, guidance, and intervention.

6-4. Develop and Improve Tools Using Big Data for Systems Science Approaches to Nutrition Research

Systems approaches, methods, and tools such as computational modeling can help to better address the system of factors and relationships that affect and are affected by nutrition. Without the use of such methods, it can be challenging to understand the complexities involved. The growth of computational capabilities continues to open new possibilities to generate novel insights from existing types of data and emerging new data types.

Systems approaches can involve the use of very large datasets or multiple data sources derived from more traditional investigations. Systems methods may provide novel conceptual opportunities to maximize the use of existing data and should help in modeling and implementation of targeted nutrition interventions. The data types and sources used with

systems approaches and methods can be diverse and complex, yielding information as varied as demographic, economic, and geographic factors; grocery purchases; proximity to grocery stores; health-related issues from electronic health records, biological variables, dietary intake, and food composition data. These datasets can be used in systems science approaches to address key questions in nutrition and to facilitate modeling of inter-relationships.

Several barriers prevent wide use of Big Data for nutrition research. For example, quantitative and real-time data, along with participants' information, are frequently maintained in databases with unique architectures, data types, and purposes, making access to and combining these datasets challenging. Overcoming these barriers would facilitate the development of comprehensive and integrated databases that could be used to address complex nutrition research questions.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Develop and support application programming interfaces (APIs) that systematically enhance researcher access to nutrition data. These applications could provide pilot data for implementation of simulated research programs informed by real-time data collection.
- Create new tools that take advantage of emerging innovations and apps to offer faster and wider (even global) participant recruitment, real-time data capture, individual data uniquely customized to a variety of aspects (e.g., genetics, sociocultural factors, environment, lifestyle practices, food intake, purchasing behavior), analysis, and data processing.
- Develop or leverage current simulation models to help represent the complex pathways that connect different systems for use in tests of various policies and interventions.
- Incorporate nutrition assessment tools into patient portals within health care delivery systems to enable comprehensive integration of nutritional, genetic, and health outcome data.

6-5. Develop Sensors for Continuous Monitoring of Nutrients and Metabolites for Personalized Nutrition Research

Continuous sensing of participants' interstitial glucose concentrations over long periods has proven valuable to identify predictive factors underlying interindividual variability in controlled feeding studies. Presently, few continuous sensors that monitor other nutrients or metabolites in

addition to glucose are available for research. Expanding the number of nutrients and metabolites that could be monitored would advance this approach to personalized nutrition research.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Leverage glucose oxidase sensor technology found in continuous glucose monitors to measure other nutrients that are currently assayed using an oxidase reaction.
- Develop and test chemical analyzers or other sensors that can be carried, placed in contact with skin, swallowed, or temporarily implanted or attached (e.g., in the oral cavity) to detect specific foods, components of foods, or metabolites.
- Miniaturize ion selective electrodes using biocompatible materials.

6-6. Develop Predictive Epigenetic Tools

Many epigenetic changes are associated with diet and nutrition. Statistical models and computer software are needed to inform predictions of when individual epigenetic modifications or patterns of modification might lead to altered gene expression. Data modeling will provide a means to incorporate inter- and intra-individual variation of diverse contextual inputs. Presently, the importance of any modification or pattern of modifications must be separately validated through transcriptomic studies.

Tissue-specific predictive tools are required because epigenetic modifications can be highly tissue-specific. Consequently, analysis of the whole blood epigenome may not always be predictive of epigenetic modifications in other tissues. Research is needed both to understand which tissues or brain regions may be responsible for epigenetic-linked phenomena such as imprinting and to determine whether those changes are faithfully replicated in blood derived DNA. Addressing that question will facilitate an understanding of what is captured and what is missed in blood-based epigenetic studies.

Future research activities that could be pursued in this Priority include, but are not limited to, the following:

- Elucidate the cell, tissue or regional specificity of epigenetic modifications responsible for individual variability in response to diet or nutrition interventions or dietary pattern interventions.
- Develop computational or modeling approaches that can help predict when patterns of epigenetic modifications are likely to affect gene expression or remain silent, and how different exposures may modify those patterns.
- Identify instances where blood epigenetic changes can be used as proxies for key tissue-specific epigenetic modifications affecting inter-individual differences in diet or nutrition responses.

6-7. Encourage the Use of Controlled Human Feeding Studies

At present, many human nutrition studies test the effects of different diet prescriptions in free-living participants. Such studies can suffer from several key limitations, including inaccurate reporting and lack of adherence to the diet prescribed. Some of these issues can be obviated by designs that use controlled feeding centers. Outcomes from controlled feeding studies are frequently rated as the highest level of evidence in meta-analyses related to nutrition. An example of where they can be useful in providing the research base to support determination of Dietary Reference Intakes, where having some certainty about the different doses of the nutrient consumed is critical. Another example is related to the development of new approaches and technologies for dietary intake assessment or improving measurement error in such measures, as proposed earlier in this Theme. Technologies and biomarkers that provide objective dietary assessment may be on the horizon but will require initial development and validation in a wide range of participants where dietary intake is known.

It is recognized that findings from controlled feeding studies will require further evaluation in free-living populations. However, they can provide the rigorous premise from which to propose larger and longer-term studies in free-living populations. Another goal of this Priority is to explore barriers and best practices in conducting research employing controlled feeding designs.

Theme 7: Support Training to Build an Outstanding Nutrition Research Workforce

Introduction

Theme 7 outlines research training priorities tailored to the needs described in the previous six Themes. Training of basic nutrition research scientists is needed to improve their skills to characterize and explain nutrition biology, physiology, and metabolism at the molecular and systems levels. Training of nutrition researchers in clinical and preclinical nutrition research is needed to help translate basic science observations into interventions that improve public health. Another critical need is for more institutions to offer combined degree training programs, particularly registered dietitian (RD)/PhD programs, to bridge the strengths of RDs' clinical experience with PhD-level training in nutrition research.

Nutrition research is a broad field of research encompassing many thought domains. Although promoting training across these domains is one way to obtain broad expertise, another approach is to provide training in how to establish and work in multidisciplinary teams to answer the broad and important questions in nutrition research. Team-based scholarship, like individual scholarship, is a learned and teachable behavior. However, only a limited number of scientists are trained in multidisciplinary, or preferably interdisciplinary, teams that include a focus on nutrition research. Thus, it is recommended that training activities provide opportunities for researchers to work with large teams, potentially early in the trainee's education. Of paramount importance for these training opportunities is to draw the best talent from a pool of candidates who are racially and ethnically diverse, as well as diverse in other areas, including sex, socioeconomic status, geographic location, and disability status.

Among the important opportunities to improve research training in nutrition research, two areas in particular include: interdisciplinary training in research related to the gut microbiome and training in utilizing Big Data.

Training Priorities

7-1. Facilitate Training in Host, Gut Microbiome Metabolism, and Diet Interrelationships

Investments in specialized training for research on host, gut microbiome, and diet interrelationships may be needed to support emerging research in this field, as described in

Priority 1-3. Training in nutritional biochemistry may be essential to enhance knowledge about the complex interactions among nutrition, biochemistry, physiology, and metabolism by the host and microbiota.

Researchers seeking to understand diet and microbiome interrelationships will require training in the biochemistry of foods, including dietary glycomics (i.e., the chemical structure and source of dietary fibers) to learn how the biochemical diversity and sources of dietary fiber shift microbial metabolism and ecology. Training gaps for research on the microbiome include integrative and GI physiology, ingestive behavior (including gut-brain communication), and the interaction between inflammation, nutritional requirements, and metabolism.

Future research training activities that could be pursued in this Priority include, but are not limited to, the following:

- Encourage training in inter-organ mammalian metabolism and comparative metabolism (eukaryotic versus microbial) and in isotopic approaches to metabolic phenotyping and metabolic flux measurements.
- Facilitate training in electrophysiological and optogenetic approaches to study microbiome-diet-host mechanisms affecting neural or enteroendocrine signaling and behavior.
- Increase training in food biochemistry and dietary glycomics and microbiota-accessible carbohydrates.

7-2. Enhance Training in the Application of Big Data Approaches to Nutrition Research

A critically important area that also demands a team-based strategy is the application of Big Data approaches to nutrition research. A new cadre of investigators needs to be trained in how to create and use quantitative methods to work with Big Data sources.

Currently, a barrier to applying Big Data methods to nutrition research is a lack of individuals trained on the integration of Big Data and systems biology. Investigators should build teams with the complete set of needed skills. The successful implementation of such studies requires multidisciplinary teams with collective expertise in: 1) model systems of host physiologic response to diet, 2) human subjects research, 3) nutritional biology, 4) high-throughput analytic technologies (e.g., metabolomics, genomics, transcriptomics, proteomics), 4) quantitative biology (e.g., computational biology, biostatistics, bioinformatics, systems biology), and 5)

biochemistry and chemical biology. To build scientific diversity, appropriate distinctions between statistical, computational, and bioinformatics expertise are needed.

Training programs in nutrition should embrace advances in digital technology to encourage future nutrition scientists to become well trained in data sciences and design-based approaches, to maximize benefits from these innovations.

Future research training activities that could be pursued in this Priority include, but are not limited to, the following:

- Increase training that fosters application of Big Data and systems biology to nutrition research.
- Facilitate the cross-training of investigators who specialize in different organ systems (including the CNS) with how to appropriately model “microbiome” to ensure that the data obtained are translatable to humans.
- Encourage nutrition research training programs to incorporate training and curricula in data sciences and Big Data approaches (or cooperative educational experiences in informatics or Big Data).
- Encourage team-based research and collaboration with scientists of diverse expertise to foster a full range of nutrition research opportunities.
- Encourage training of clinicians in translational research and implementation research related to nutrition across the lifespan to improve patient health and well-being.

CONCLUSION

The pursuit of research opportunities related to the Themes and Priorities identified in this *Strategic Plan* will help accelerate progress in improving health. As the Nutrition Research Task Force transitions into the implementation phase of this *Strategic Plan*, many NIH program staff dedicated to the advancement of nutrition science will meet regularly to monitor progress and to coordinate and discuss opportunities to advance the priorities identified in each of the seven Themes and cross-cutting areas.

Investigator-initiated efforts are expected to serve as the major driver of progress in nutrition research, as is the case across biomedical research supported by the NIH. The Task Force will use the Themes and Priorities identified in this *Strategic Plan* to monitor progress through portfolio analyses, scientific publications, and workshops and will catalyze extramural investigations through NIH-solicited research programs and new Funding Opportunity Announcements. Throughout the next 10 years, the broad nutrition research community will also be engaged through workshops and other processes to further refine specific activities that will be employed to advance the priorities in this *Strategic Plan*. These efforts will enable a deeper understanding of the interactions between diet, nutritional status, biological processes, and the environment, leading to new approaches and interventions to prevent and treat disease and reduce health inequity. To ensure success of this *Strategic Plan* in advancing future nutrition research, priorities may evolve in response to emerging opportunities and a changing scientific landscape.