Human Microbiome Project – Core Microbiome Sampling Protocol A

HMP Protocol Number:
HMP-07-001

Sponsored by:
National Institutes of Health (NIH)

NIH Funding Mechanism:
U54-HG003273 – Baylor College of Medicine
U54-HG003079 – Washington University
N01-HG-62088 - Coriell Institute of Medical Research

Draft or Version Number:
Version 9.0

Date:
22 Mar 2010

Principal Lead Sites:

<table>
<tr><th colspan="2">Coordinating Investigator and Study Chair</th></tr>
<tr><td colspan="2">James Versalovic, MD, PhD
Baylor College of Medicine
Texas Children’s Hospital
6621 Fannin Street, MC 1-2261
Houston, TX 77030
(832) 824-2213
jamesv@bcm.edu</td></tr>
<tr><th colspan="2">Baylor College of Medicine</th></tr>
<tr><td>Wendy Keitel, MD
Dept. of Molecular Virology and Microbiology
Baylor College of Medicine 280
One Baylor Plaza
Houston, TX 77030
(713) 798-5250
wkeitel@bcm.edu</td><td>Joseph Petrosino, PhD
Baylor College of Medicine
One Baylor Plaza, N1519, MS: 226
Houston, TX 77030
(713) 798-7719
jpetrosi@bcm.tmc.edu</td></tr>
<tr><th colspan="2">Washington University</th></tr>
<tr><td>Mark Watson, MD, PhD
Washington University School of Medicine
Campus Box 8118
660 S. Euclid Avenue
St. Louis, MO 63110
(314) 454-7919
watsonm@wustl.edu</td><td>Michael Dunne, PhD
Washington University School of Medicine
258A Barnes-Jewish Hospital Service Building
St. Louis, MO 63110
(314) 362-1547
dunne@wustl.edu</td></tr>
</table>

NIH Contact Persons:

Pamela McInnes, DDS, MSc.(Dent.)
Director, Division of Extramural Research
National Institute of Dental and Craniofacial Research (NIDCR)
National Institutes of Health
6701 Democracy Blvd.
Room 610, Mail Stop 4878
Bethesda, MD 20892-4878
(301) 443-8618
pmcinnes@nidcr.nih.gov

Jane Peterson, PhD
Associate Director, Division of Extramural Research
National Human Genome Research Institute (NHGRI)
National Institutes of Health
5635 Fishers Lane Suite 4076
Bethesda, MD 20892-9305
(301) 496-7531
petersoj@exchange.nih.gov

Jean E. McEwen, JD, PhD
Program Director, Ethical, Legal, and Social Implications Program
National Human Genome Research Institute (NHGRI)
National Institutes of Health
5635 Fishers Lane, Suite 4076, MSC 9305
Bethesda, MD 20892-9305
(301) 496-7531
jean.mcewen@nih.hhs.gov

Michelle A. Culp, BSN, MPH
Director, Office of Clinical Trials Operations and Management
National Institute of Dental and Craniofacial Research (NIDCR)
National Institutes of Health
Democracy I, Room 644
Bethesda, MD 20892-4878
(301) 594-4830
mculp@nidcr.nih.gov

Coordinating Authors:

Pamela McInnes, DDS, MSc.(Dent.)
Mary Cutting, MS, RAC, NIH Contractor

Statement of Compliance

The clinical study will be carried out in accordance with Good Clinical Practice (GCP) as required by the following documents:

- U.S. Code of Federal Regulations applicable to clinical studies (45 CFR 46)
- ICH GCP E6
- Completion of Human Subjects Protection Training
- NIH Clinical Terms of Award

Refer to: *http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.htm#46.*
http://www.fda.gov/cder/guidance/959fnl.pdf
http://grants.nih.gov/grants/guide/notice-files/NOT-OD-01-061.html
http://cme.cancer.gov/c01/

Signature Page

I have read the protocol, including all of the appendices and the Manual of Procedures. I agree that these documents contain all of the necessary information to conduct this study as described. I will conduct this study as outlined herein, in accordance with the regulations stated in the Federal Code of Regulations for Protection of Human Subjects Title 45 and the International Conference on Harmonisation Good Clinical Practices Guidelines (E6), and will make a reasonable effort to complete the study within the time designated.

Clinical Site Principal Investigator:

Signed: Date:

_____ _____

Title

Table of Contents

page

List of Abbreviations

AE	Adverse Event
CDCC	Clinical Data Coordinating Center
CFR	Code of Federal Regulations
CLIA	Clinical Laboratory Improvement Amendments
CoC	Certificate of Confidentiality
CSOC	Clinical Study Oversight Committee
DHHS	Department of Health and Human Services
DNA	Deoxyribonucleic acid
FWA	Federal-Wide Assurance
GCP	Good Clinical Practice
GI	Gastrointestinal
HBV	Hepatitis B virus
HCV	Hepatitis C virus
HIPAA	Health Insurance Portability and Accountability Act
HIV	Human immunodeficiency virus
HMP	Human Microbiome Project
HPV	Human papillomavirus
IBD	Inflammatory bowel disease
IBS	Irritable bowel syndrome
ICH	International Conference on Harmonisation
IDES	Internet Data Entry System
IRB	Institutional Review Board
JAMA	Journal of the American Medical Association
MOP	Manual of Procedures
N	Number (typically refers to subjects)
NEJM	New England Journal of Medicine
NHGRI	National Human Genome Research Institute
NIDCR	National Institute of Dental and Craniofacial Research
NIH	National Institutes of Health
OHRP	Office for Human Research Protections
PD	Protocol Deviation
PHI	Protected Health Information
PI	Principal Investigator
rRNA	Ribosomal ribonucleic acid
SOP	Standard Operating Procedure
STD	Sexually transmitted disease
VFH	Volunteer for Health

Glossary of Terms

Subject(s): Term used throughout the protocol to denote an individual who is contacted to become a participant, who is screened or who is enrolled in the study

Eligible: Qualified for enrollment into the study based upon strict adherence to inclusion/exclusion criteria

Microbial nucleic acids: DNA and RNA from microorganisms including bacteria, viruses, fungi and protozoa

Minority population: A human subjects term indicating a subset of the U.S. population distinguished by racial, ethnic, or cultural heritage

Protocol administrative change: A change to a protocol documented in a memo informing the IRB of a change(s) to only logistical or administrative aspects of the study

Protocol amendment: A written description of a change(s) to or formal clarification of a protocol, prepared as an amendment to the protocol; any change to an approved clinical protocol that affects the safety of subjects, the scope of the investigation, the design of the study, or scientific integrity of the study (e.g., the collection of specimens in an invasive manner that represents more than minimal risk)

Screened Subject: Subject who has provided consent for participation and has been screened for the research; subject may or may not be eligible

Study monitor: An individual centrally located at NIH who is responsible for monitoring proper conduct of the clinical study

Site monitor: An individual assigned by NIH who is responsible for monitoring conduct at individual study sites and providing feedback to investigators and the NIH

Protocol Summary

Title: **Human Microbiome Project – Core Microbiome Sampling**

Population: Approximately 300 male and female adults, 18-40 years of age, in good health status, from two geographic locations in the US.

Number of Sites: Two sites: Baylor College of Medicine, Houston, TX and Washington University, St. Louis, MO

Study Duration: Approximately 18 months

Duration of Subject's Participation: A few weeks up to 18 months

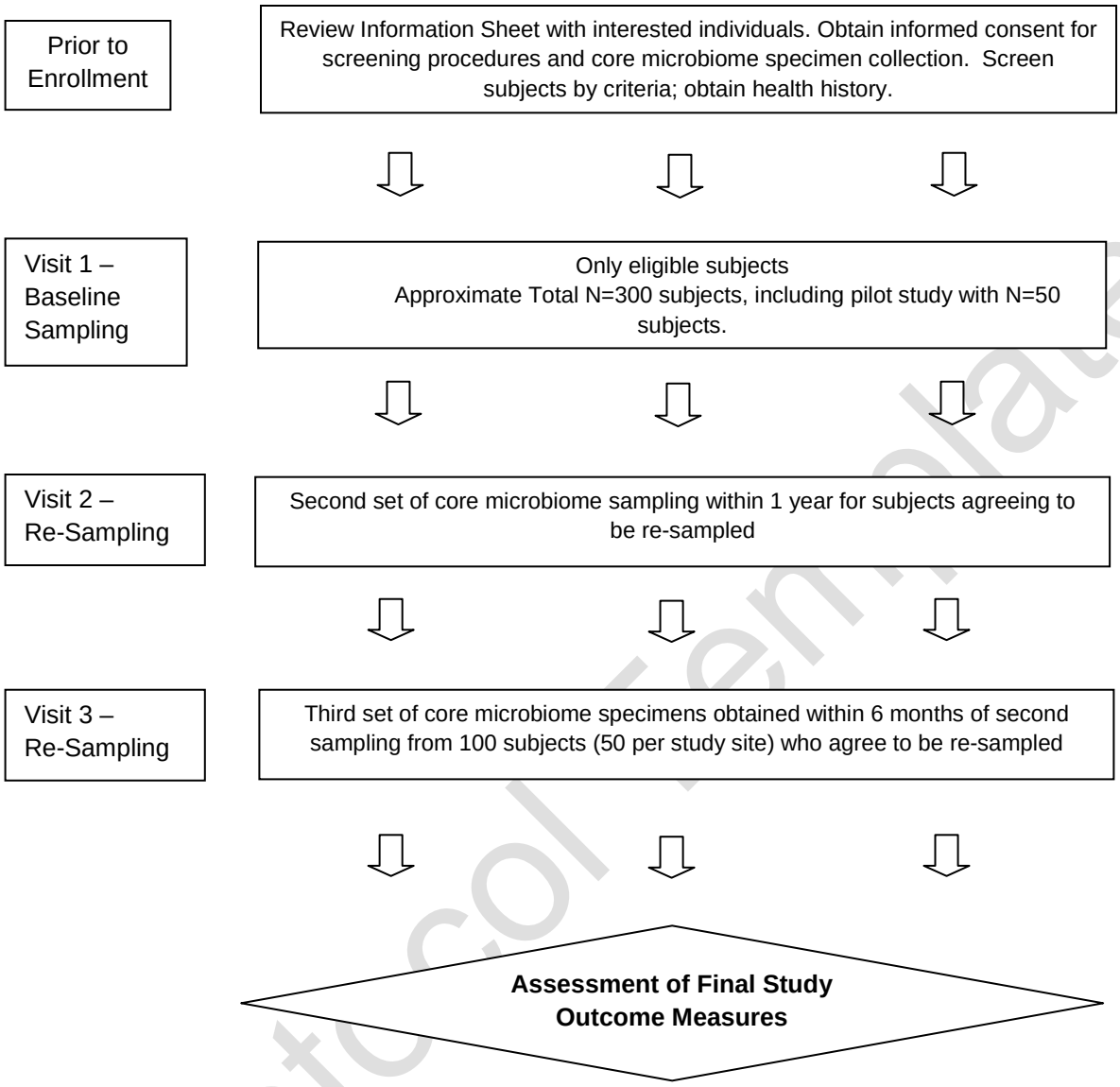
Objectives:

- To collect clinical specimens from multiple body sites in a non-invasive manner to serve as sources of measurement of the core microbiomes associated with the oral cavity, skin, nasal cavity, gastrointestinal tract and vagina.

- To collect blood from the subjects donating clinical specimens, to be used to examine the relationship of host genotype (by either genotyping or sequencing the DNA) to the microbiota present on/in an individual.

- To collect serum for possible future research studies, such as determinations of immune responses to organisms that are identified in the microbiome.

Schematic of Study Design



1 KEY ROLES

For questions regarding this protocol, contact:

NIH Representative:

Pamela McInnes, DDS, MSc(Dent.)
Director, Division of Extramural Research
NIDCR, NIH
6701 Democracy Blvd.
Room 610, Mail Stop 4878
Bethesda, MD 20892-4878
(301) 443-8618
(301) 480-8319 (fax)
pm23v@nih.gov

Coordinating Investigator and Study Chair:

James Versalovic, MD, PhD
Professor of Pathology
Baylor College of Medicine
Texas Children's Hospital
6621 Fannin Street, MC 1-2261
Houston, TX 77030
(832) 824-2213
jamesv@bcm.edu

Baylor College of Medicine Principal Investigator:

Wendy Keitel, MD
Professor
Department of Molecular Virology and Microbiology
Baylor College of Medicine 280
One Baylor Plaza
Houston, TX 77030
(713) 798-5250
wkeitel@bcm.edu

Baylor College of Medicine Co-Investigator:

Joseph Petrosino, PhD
Baylor College of Medicine
One Baylor Plaza, N1519, MS: 226
Houston, TX 77030
(713) 798-7719
jpetrosi@bcm.tmc.edu

Washington University Principal Investigator:

Mark Watson, MD, PhD

Washington University School of Medicine
Campus Box 8118
660 S. Euclid Avenue
St. Louis, MO 63110
(314) 454-7919
watsonm@wustl.edu

Washington University Co-Investigator:

Michael Dunne, PhD
Washington University School of Medicine
258A Barnes-Jewish Hospital Service Building
St. Louis, MO 63110
(314) 362-1547 (office)
(314) 362-1306 (lab)
dunne@wustl.edu

Clinical Data Coordinating Center:

Gina A. Gorgone Simone
The EMMES Corporation
401 N. Washington St., Suite 700
Rockville, MD 20850
(301) 251-1161 (main)
(301) 251-1161 x2810 (direct)
(301) 251-1355 (fax)
gsimone@emmes.com

Blood Specimen Repository:

Neda Gharani, PhD
Coriell Institute for Medical Research
NHGRI Sample Repository for Human Genetic Research
403 Haddon Ave.
Camden, NJ 08103
(856) 757-9722
(856) 757-9737 (fax)
ngharani@coriell.org

2 BACKGROUND INFORMATION AND SCIENTIFIC RATIONALE

2.1 Background Information

The collections of microbes that reside in the human body form complex communities and are found primarily in the oral cavity, on the skin, in the nasal cavity, in the gastrointestinal tract, and in the vagina (Paustian 2006; Tierno 2001). These microbial communities are intimately integrated with the human "self" and may lead to a new understanding of "self". They vary with many factors, such as age, gender, environment, diet and disease state. How these variations are accomplished and the impact of these variations on human development, physiology, immunity, and nutrition is not understood. To pursue this understanding by taking advantage of recent technological advances, the NIH Roadmap has initiated the Human Microbiome Project (HMP), with the mission of generating resources to enable comprehensive characterization of the human microbiota and analysis of their role in human health and disease.

Traditional microbiology has focused on the study of individual species as isolated units. However many, if not most, have never successfully been isolated as viable specimens for analysis, presumably because their growth is dependent upon a specific microenvironment that has not, or cannot, be reproduced experimentally. Among those species that have been isolated, analyses of genetic makeup, gene expression patterns, and metabolic physiologies have rarely extended to inter-species interactions or microbe-host interactions. Advances in DNA sequencing technologies have created a new field of research, called metagenomics, allowing comprehensive examination of microbial communities, even those comprised of uncultivable organisms. Instead of examining the genome of an individual bacterial strain that has been grown in a laboratory, the metagenomic approach allows analysis of genetic material derived from complete microbial communities harvested from natural environments. In the HMP, this method will complement genetic analyses of known isolated strains, providing unprecedented information about the complexity of human microbial communities.

By leveraging both the metagenomic and traditional approaches to genomic DNA sequencing, the HMP will lay the foundation for further studies of human-associated microbial communities. Broadly, the project has set the following goals:

- Determining whether individuals share a core human microbiome

- Understanding whether changes in the human microbiome can be correlated with changes in human health
- Developing the new technological and bioinformatic tools needed to support these goals
- Addressing the ethical, legal and social implications raised by human microbiome research.

This first clinical study addresses the first HMP goal of determining whether individuals share a core human microbiome.

2.2 Scientific Rationale

This study involves broad determination of the microbiota found in five anatomical sites: the oral cavity, skin, nasal cavity, gastrointestinal tract and vagina. This study will enroll healthy adults from two geographic regions of the US. The participation of healthy individuals will create a baseline for discovery of the core microbiota typically found in various areas of the human body. The information from this initial study can then be used to help assess the changes in the complement of microbiota found on or within diseased individuals. The specimens will be collected from approximately equal numbers of males and females. All subjects will be between the ages of 18 and 40. The age range of subjects has been established to minimize the influence of growth and developmental changes as well as those changes associated with the aging process.

2.3 Potential Risks and Benefits

2.3.1 Potential Risks

The methods of specimen collection in this first core microbiome sampling study protocol pose only minimal risk to the study subjects. As defined in 45 US Code of Federal Regulations (CFR) 46.102 (i), “Minimal risk means that the probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests.”

Neither immediate nor long-range physical risks are involved. There are no alternative procedures or protocols available for conducting this research. The information gained from this study will help to evaluate later studies conducted with diseased individuals. Any gain made in studying these conditions would outweigh the risk to individuals who participate in this study.

Minimal physical risks are listed below:

- Blood specimen: There may be brief discomfort or pain and bruising at the site of venipuncture. There is also a very small chance that an infection could result or that there could be excessive bleeding. The subject may become dizzy or faint from the blood draw.
- Oral cavity: The subject could experience a dry area in the mouth at the sampling site for a short period of time (less than 5 minutes). Slight tenderness could occur in the gingivae (gums) surrounding teeth where specimens have been collected.
- Skin: Rubbing of the skin may cause slight irritation or transient redness at the sampling sites.
- Nasal cavity: There may be slight irritation at the sampling sites.
- Stool specimen: Contamination of skin with feces from the collection container could occur.
- Vagina: There may be slight irritation at the sampling sites and some subjects may experience embarrassment at having the procedure.

The nature of the information collected from subjects may involve risk to their privacy. This study will create a resource that will be available over the internet. The microbial data, which will be coded, will be deposited and be available through an open-access database. The clinical data and the eventual counterpart human sequence data, which will also be coded, will be available through a controlled-access database, accessible only to qualified investigators through an NIH Data Access Committee. DNA (and possibly later lymphoblastoid cell lines) from the blood specimens, which will be coded, may be made available to investigators outside the project for other studies of the human microbiome; however, the repository that distributes the specimens will not have access to the codes. If the subject grants permission, coded serum specimens held at the clinical site repositories may be distributed to other researchers for future studies not necessarily related to the human microbiome. Appendices A and B contain tables and a flowchart that address the disposition and access issues for original specimens as well as the DNA samples, possible cell lines, serum specimens and sequence data that may be derived from the original specimens.

There is a small risk that some information could be disclosed to someone outside of the project or that, at some point, information stored in the controlled access databases could in some way be linked back to a specific subject. Despite the recent passage of a federal law that bars genetic discrimination in employment and some types of insurance, the use by others of genetic information from the eventual study of the blood specimens could conceivably impact the subject negatively in some way. For example,

there is a very small risk that genetic information obtained by this study could be used by law enforcement officials to try to learn more about the subjects or their family members for the purpose of a criminal investigation. However, a Certificate of Confidentiality (CoC) will be obtained that will prevent the participating institutions from being forced to disclose identifying information about subjects, without their permission, for use in such cases. In addition, anyone seeking access to the information in the controlled access databases will need prior approval of an NIH Data Access Committee. The controlled access databases are intended for research use only and the NIH Data Access Committee will not approve other uses such as for criminal investigations.

The results of HIV, hepatitis B, and hepatitis C tests will be available from the screening process. The results of pregnancy screening in women will also be available. These results will be provided to the subjects. Subjects with positive results will be referred for follow-up medical care. Positive results are exclusion criteria for the study so the subject will not be eligible to be enrolled. An unanticipated positive result from these tests has the potential to have adverse psychological impact on the screened subject.

Microbial nucleic acids (DNA and RNA) from pathogens reportable to public health authorities (e.g., reportable sexually transmissible pathogens), and from other pathogens with potential clinical significance may be discovered during the process of human microbiome studies. This information will not be shared with the subject, the subject's physician or the state public health authorities because the sequence data will be obtained in a non-CLIA certified setting, and will therefore not be validated for clinical diagnosis. Furthermore, the plan to make both the nucleic acid samples and the sequence databases available (with all information coded) for further study by numerous investigators makes it impracticable to commit to reporting of clinically significant findings to individual subjects or to their physicians. If investigators become aware of data with exceptional potential clinical significance, they will determine how best to proceed, in consultation with the NIH and ethics advisors. Information on this issue is available at http://bioethics.od.nih.gov/internationalresthics.html.

It is possible that the microbial nucleic acid specimens may contain small amounts of human DNA or RNA. Every effort will be made to remove any traces of human DNA or RNA sequence data from the microbial data by computational or other means, but it is possible that a very small amount of human nucleic acid (DNA or RNA) sequence data will remain.

The results of research cannot be foreseen so it is possible that new risks may arise in the future that cannot be predicted now. However, it is believed that the benefits of learning more about the human microbiome and how it relates to health and disease outweigh the current and potential future risks.

2.3.2 Known Potential Benefits

Subjects who are screened for eligibility will be compensated for their time and effort. Eligible subjects enrolled in the study will be reimbursed for completion of a study visit, including collection of body site specimens. Other than this, subjects will not derive direct benefit of either money or treatment for participation. They will not benefit personally from giving the specimens because this research will take a long time to produce medically useful results. However, the subjects will be part of a groundbreaking biological study that will help researchers around the world and the general public gain knowledge and understand more about the human microbiome and how it relates to health and disease. A potential benefit of enrolling in the study is that subjects will receive a medical evaluation by health care professionals.

3 OBJECTIVES

The primary objective of this study is to collect specimens from multiple body sites of approximately 300 healthy human adults in a non-invasive manner to serve as sources of measurement of the core microbiomes associated with the oral cavity, skin, nasal cavity, gastrointestinal tract, and vagina. Blood specimens will be collected from these same subjects and will be used to create repositories of human DNA for future sequencing and cryopreserved lymphocytes for possible creation of lymphoblastoid cell lines. Serum specimen repositories will also be created for possible future measurement of serum immune responses.

The study will have an initial pilot phase in which 50 eligible subjects are enrolled, in order to validate collection procedures and logistics. Approximately 250 additional eligible subjects will then be enrolled and will have specimens collected from all or some of the same body sites as the pilot group. Study subjects will be asked to return within approximately a 1 year period for a second round of sampling; subjects will be asked to reconfirm their consent at the time they return. The collection event alone will test the ability of investigators to effectively collect microbiome specimens across multiple sites and individuals and isolate nucleic acids suitable for sequencing.

Preliminary analysis of data from this study suggests that collection of a third set of specimens could provide important information about the microbial diversity within an individual over time. To obtain this information, approximately 100 subjects (50 from each of the two study sites) will be asked to return within 6 months of the second sampling for a third sampling of all body sites. Study subjects who return for a third sampling will be asked to sign a new informed consent at the time of that sampling. A subsequent protocol (Core Microbiome Specimen Protocol B) may undertake more invasive sampling of the gastrointestinal tract in a small subset of subjects (approximately 10).

Following specimen collection, coding and processing, the sequencing of the isolated microbial nucleic acids will produce a list of core microbiota found in each location. In total, the data produced from this project will eclipse any known set of data and provide a core microbiome catalogue. These data will also allow more detailed and significant conclusions about inter- and intra-subject diversity as well as the overall composition of the human microbiome and how it changes over time. Producing this catalogue will require the development of novel bioinformatics tools used to align the data to reference genome sequences and an analysis pipeline designed to model the interaction of these bacteria in the environment from which they were collected.

The data sets from these studies will be considered a community resource to be used for further study. Methods of data storage, display and analysis will be developed to enable the community at large to take advantage of the data to the fullest extent. This analysis will then be used to direct new avenues of research for both sampling research and disease-based research. Comparisons of information derived from these specimens will be used to catalogue the microbiota that exist in the general population and the variation that exists within them to determine whether individuals do indeed share a core microbiome.

The whole blood specimens collected will be coded and sent to the NHGRI Sample Repository for Human Genetic Research at Coriell Institute, which will extract human DNA from them and possibly, in the future, make lymphoblastoid cell lines. The repository will store and maintain the coded human DNA and/or cell lines and will distribute these materials for future approved studies related to the human microbiome. Serum samples from the study subjects will be stored in repositories at the clinical research sites for possible use in future research, such as assaying immune responses to organisms that are identified in the microbiome.

It is anticipated that the specimens collected from the oral cavity, skin, nasal cavity, and vagina will be completely consumed in the DNA extraction process. The quantity of material collected in the stool specimen will likely be more than is required for DNA extraction. The primary focus of this study is extraction and sequencing of microbial DNA. If the volume of clinical specimen allows, it is possible that RNA extraction may be undertaken from the same specimens for RNA sequencing and gene expression studies. Residual body site specimens will be retained for a period of time to permit the application of secondary and/or alternative nucleic acid extraction methodologies. After these techniques have been implemented and the sequencing is completed, remaining body site specimens will be destroyed. In order to complete metagenomic sequencing, it may be necessary to amplify the isolated microbial DNA. Microbial nucleic acids remaining after sequencing is completed may be made available to other researchers for future approved studies related to the human microbiome.

4 STUDY DESIGN OVERVIEW

This study is designed to collect specimens from healthy adults from five body sites: the oral cavity, skin, nasal cavity, gastrointestinal tract, and vagina (female subjects only). Originally, the total enrollment target for the study was 250 subjects. Preliminary data analyses have yielded information regarding inter-individual and intra-individual variation in the composition of the microbiome. The extent of intra-individual variation between visits 1 and 2 is body site-dependent, and has generated interest in additional sampling efforts. As the study approaches the successful recruitment of 250 healthy, adult subjects in St. Louis and Houston, the study Steering Committee has decided to expand the study in two basic ways. First, additional individuals (25 per site) will be enrolled, so that a total of 300 subjects will be sampled. This enhancement represents an increase of 20 percent in total subjects sampled, so that compositional data from multiple body sites are even more robust. Secondly, the study will add a third sampling of 100 subjects (50 per site). A third sampling will enable investigators to study intra-individual variation during an extended time period, and intra-individual variation among different body sites can be analyzed more comprehensively.

More than 300 interested individuals will need to be screened in order to derive approximately 300 eligible subjects. The study will be phased, with 50 eligible subjects participating in an initial pilot phase and having four or five body sites sampled as well as having a blood specimen collected. Following evaluation of the operational aspects of the pilot phase, approximately 250 additional eligible subjects will be enrolled and these subjects will have up to five of the same body sites sampled and will have a blood specimen drawn.

Individuals involved in the study will be between the ages of 18 and 40 and sampling will include an approximately equal number of males and females, with a goal of at least 20% representation from minority populations. Subjects will be contacted to have the same body sites sampled a second time within approximately 1 year of the first sampling. Approximately 100 subjects (50 subjects from each of the two sites) will be contacted to have the same body sites sampled a third time within approximately 6 months of the second sampling.

Subjects will be screened using a general health questionnaire and will be examined by a health professional to verify eligibility.

Eligible and consented subjects will be enrolled and specimens will be collected over the course of approximately eighteen months. The duration of individual subject

participation will range from a few weeks to approximately eighteen months (for subjects who provide a third set of specimens). Each subject will provide at least one set of specimens and all subjects will be asked to provide a second set of specimens within a one year time frame (although it is anticipated that at least some subjects, after being re-contacted, will be unable or unwilling to return for re-sampling). Study subjects who provide a second set of specimens will be informed that a third set of specimens may be requested, and will be asked whether they agree to be contacted for a possible third sampling. Approximately 100 subjects (50 per site) will be sampled a third time. Re-sampling will allow different collection time points to be compared and assessed for differences. In addition to this, ten (10) of the approximately 300 subjects may be invited back at a future date for more invasive sampling of the gastrointestinal tract under a separate study protocol.

Specimen collection methods will include phlebotomy by a medical professional to obtain a blood specimen, self-collection of stool specimens, swabbing by trained medical professionals to obtain collections from the mouth, skin surfaces, nose, and in female subjects, the vagina. Specific subject information will be collected via conversational interviews, with information being recorded on health history forms and clinical report forms.

After specimen collection and coding, microbial DNA will be isolated from the oral cavity, skin, nasal cavity, gastrointestinal tract and vaginal specimens. The DNA will then be used to produce DNA sequence information. Sequence data will be generated using a variety of platforms that are used to read the sequence of chemicals that make up the structure of DNA found in these specimens. The whole blood specimens will be processed for human DNA extraction and cryopreservation of lymphocytes. At a later time, genotyping and/or sequencing will be undertaken using the DNA samples and lymphoblastoid cell lines will likely be derived from the cryopreserved lymphocytes. Serum samples will be stored with no immediate analysis performed; these may be used later for other research studies, such as measuring immune responses to organisms that are identified in the microbiome.

The primary outcomes of this project are the collection of the aforementioned specimens, isolation of DNA from them, and subsequent use of the DNA to produce sequence data that represent the microbiota found in each specimen (or, in the case of the blood specimens, human DNA sequence data). Secondary outcomes include the establishment of serum repositories and repositories for DNA and (possibly) lymphoblastoid cell lines, and development of tools for analysis. Together the information obtained in this study will provide a reference catalogue of information, tools for analysis and data resources for the research community.

5 STUDY POPULATION

5.1 Selection of the Study Population

The target for this multi-center study is approximately 300 healthy subjects between the ages of 18 and 40. Approximately half of the subjects will be female and approximately half will be male. A goal for minority participation will be at least 20% of the total number of subjects.

At Baylor College of Medicine (BCM), subjects will be recruited through the Vaccine Research Center under the direction of Wendy Keitel, MD. Dr. Keitel will serve as Clinical Principal Investigator for the study and will supervise the nursing team trained in clinical assessment (histories/physicals) and human sampling. Healthy adults will be recruited from the Texas Medical Center (TMC) and environs. The TMC represents over 42 member institutions, which includes 13 hospitals and 2 specialized patient facilities. Over 65,300 people were employed in the TMC in 2004. Eleven educational institutions are located in the TMC or are affiliated with the TMC; these enroll over 22,000 students. College students at Rice University (enrollment~5,000), the University of Houston (enrollment = 35,180), and other educational institutions also can be recruited for participation. Healthy adults also can be recruited from surrounding industrial populations such as oil companies, as has been done in the past. In addition, the BCM Vaccine Research Center maintains a registry of people who are interested in being contacted regarding upcoming studies, and the Office of Research at BCM maintains a Volunteer Line and registry that lists people who have consented to be contacted by an investigator from BCM for possible enrollment in an IRB approved protocol. Potential subjects will be recruited by posting fliers around the TMC, by sending email announcements via approved listserv distribution lists, and by contacting individuals who have given permission via email or telephone.

Facilities at Baylor College of Medicine will enable all sampling procedures to occur at a single site. Oral assessment and sampling will be performed with the assistance of Catherine Flaitz, DDS, and her team of dental health care professionals including dental hygienists from the University of Texas Health Sciences Center – Dental Branch. Vaginal sampling will be performed with the assistance of Kjersti Aagard, MD, PhD, Amy Young, MD and the team of health care professionals from the Department of Obstetrics and Gynecology at Baylor College of Medicine. Joseph Petrosino, PhD will serve as co-investigator responsible for the DNA extraction laboratory. James Versalovic, MD, PhD will serve as Coordinating Investigator and Study Chair and will

supervise a team of laboratory professionals from Baylor College of Medicine involved in specimen collection, processing, storage, and distribution.

At Washington University, Dr. Mark Watson is a member of the research faculty and as such, has access to the services offered by the Clinical Trials Unit. The Clinical Trials Unit will primarily recruit subjects through the Volunteer for Health (VFH) program. The Volunteer for Health program is sponsored by Washington University School of Medicine to support patient-oriented research at Washington University Medical Center. The overall purpose of VFH is to match interested volunteers with current clinical studies at the medical school. Currently, VFH has a database of 25,000+ prospective volunteers (persons who provided VFH with general health information and who gave their consent to be contacted about studies) that can be searched based on the initial inclusion/exclusion study criteria. Additionally they will place IRB-approved fliers, posters and advertisements in appropriate locations and media outlets in order to reach out to potential participants. Working with the VFH program will greatly facilitate tracking of subjects for resampling during this study and for participation in future studies related to the human microbiome.

Dr. Watson will serve as the Clinical Principal Investigator for the study and will oversee the collection of samples and generation of secondary resources. Dr. Watson is also Director, Siteman Cancer Center Tissue Procurement Core Facility. This entity will participate in purifying microbial nucleic acids as well as in storing and distributing the microbial nucleic acid samples. In this way, Dr. Watson will have complete oversight of the entire sample collection process. Sally Anderson, who serves as the Director of Clinical Research Services, will directly supervise the clinical nursing team collecting the samples. Dr. Mark Watson and Dr. Michael Dunne are leading a group of WU investigators who will provide ongoing support and oversight of this study. In addition to these professionals, Nathalia Garcia, DDS, will collect the oral samples and Tessa Madden, MD, will collect the vaginal samples.

5.2 Inclusion/Exclusion Criteria

A pre-screening questionnaire will be utilized prior to scheduling the Screening Visit in order to screen out potential ineligible subjects by history.

Inclusion Criteria:

In order to be eligible for participation in this study, subjects must meet the following criteria:

- Male or female subjects 18 years of age, but not more than 40 years of age at the time of enrollment.

- Must be able to provide signed and dated informed consent.

- Healthy subjects willing and able to provide blood, as well as oral cavity, skin, nasal cavity and stool specimens; female subjects must be willing to provide a vaginal specimen and must either have regular menstrual cycles (between 21 and 35 days) or, for subjects on hormonal contraception influencing cycle length, have a history of regular 21 to 35 day menstrual cycles prior to initiating hormonal contraception. At study enrollment, female subjects may be using any contraception method except a combination hormone vaginal ring (see Exclusion Criteria). (Section 6.2, Detailed Description of Study Procedures, includes information on other vaginal contraceptive products that must be avoided during the 48 hours prior to sampling.)

Exclusion Criteria:

Any subject who meets any of the following criteria will be excluded from participation in this study:

- Body Mass Index greater than or equal to 35 or less than or equal to 18.

- Vital signs outside of acceptable range at Screening Visit, i.e., blood pressure >160/100, oral temperature >100°F, pulse >100.

- Use of any of the following drugs within the last 6 months:

    - systemic antibiotics, antifungals, antivirals or antiparasitics (intravenous, intramuscular, or oral);
    - oral, intravenous, intramuscular, nasal or inhaled corticosteroids;
    - cytokines;
    - methotrexate or immunosuppressive cytotoxic agents;
    - large doses of commercial probiotics consumed (greater than or equal to 10^8 cfu or organisms per day) - includes tablets, capsules, lozenges, chewing gum or powders in which probiotic is a primary component. Ordinary dietary components such as fermented beverages/milks, yogurts, foods do not apply.
    - for female subjects, combination hormone vaginal ring for contraception (due to unknown duration of local hormone effects).

- Receipt of nasally-delivered live, attenuated, cold-adapted influenza vaccine within the previous 28 days.

- Use of topical antibiotics or topical steroids on the face, scalp, or neck or on arms, forearms, or hands within the previous 7 days.

- Use of vaginal/vulvar medications, including antifungals, within the previous 7 days. Subjects may continue to use permitted vaginal contraceptives until 48 hours prior to sampling.

- Acute disease at the time of enrollment (defer sampling until subject recovers). Acute disease is defined as the presence of a moderate or severe illness with or without fever.

- Chronic, clinically significant (unresolved, requiring on-going medical management or medication) pulmonary, cardiovascular, gastrointestinal, hepatic or renal functional abnormality, as determined by medical history or physical examination.

- History of cancer except for squamous or basal cell carcinomas of the skin that have been medically managed by local excision.

- Unstable dietary history as defined by major changes in diet during the previous month, where the subject has eliminated or significantly increased a major food group in the diet.

- Recent history of chronic alcohol consumption defined as more than five 1.5-ounce servings of 80 proof distilled spirits, five 12-ounce servings of beer or five 5-ounce servings of wine per day.

- Positive test for HIV, HBV or HCV.

- Any confirmed or suspected condition/state of immunosuppression or immunodeficiency (primary or acquired) including HIV infection.

- Major surgery of the GI tract, with the exception of cholecystectomy and appendectomy, in the past five years. Any major bowel resection at any time.

- History of active uncontrolled gastrointestinal disorders or diseases including:
    - inflammatory bowel disease (IBD) including ulcerative colitis (mild-moderate-severe), Crohn's disease (mild-moderate-severe), or indeterminate colitis;

    - irritable bowel syndrome (IBS) (moderate-severe);
    - persistent, infectious gastroenteritis, colitis or gastritis, persistent or chronic diarrhea of unknown etiology, *Clostridium difficile* infection (recurrent) or *Helicobacter pylori* infection (untreated);
    - chronic constipation.

- Regular urinary incontinence necessitating use of incontinence protection garments.
- Female who is pregnant or lactating.
- Condyloma or human papillomavirus (HPV) diagnosis within the previous 2 years.
- Treatment for or suspicion of ever having had toxic shock syndrome.
- For females, history of candidiasis, urinary tract infection, or active STD (specifically chlamydia, gonorrhea, syphilis, genital herpes, trichomoniasis) within the previous 2 months.
- For females, history of vulvar, vaginal or cervical dysplasia within the previous 5 years.
- History of hysterectomy.
- Vaginal pH greater than 4.5 at screening visit.
- Evidence of vulvar or vaginal irritation at screening or on specimen collection day.
- History of psoriasis or recurrent eczema. Childhood eczema that has resolved is not exclusionary.
- History of recurrent rashes within the past 6 months.
- At the time of the screening visit or on the specimen collection day:
    - acne at sites other than on the face, chest, back or shoulders;
    - multiple blisters, pustules, boils, abscesses, erosions or ulcers on the scalp, face, neck, arms, forearms or hands;
    - a single blister, pustule, boil, abscess, erosion, ulcer, scab, cut, crack or pink/hyperpigmented patch or plaque at or within 4 cm of the sampling

sites; sampling may be deferred until the lesion resolves either without treatment or with local treatment only;

- more than one pink/red scaly patch/plaque anywhere on the body (suggestive of psoriasis or eczema);
- uniformly thickened, cracking, “dry” skin on bilateral palms and/or soles;
- scalp dandruff that does not clear up with over-the-counter dandruff shampoos used daily for 2 weeks;
- disseminated rash (at multiple body sites or extending throughout a broad body area).

- Chronic dry mouth, as assessed through questioning of the subject by an experienced clinician.
- Periodontal pockets equal to or greater than 4 mm. (Mild gingivitis is acceptable.)
- More than 10% of sites with bleeding on probing.
- Evidence of untreated cavitated carious lesions or oral abscesses.
- Evidence of precancerous or cancerous oral lesions.
- Evidence of oral candidiasis.
- Evidence of halitosis, as determined by organoleptic assessment by an experienced clinician.
- More than 8 missing teeth. The missing teeth must be due to 3rd molar extractions and/or teeth extracted for orthodontic purposes, teeth extracted as a result of trauma, or teeth that are congenitally missing.

6 STUDY PROCEDURES/EVALUATIONS

6.1 Outline of Study Procedures

Visit Timing Sampling Timepoint		Screening Days -30 to -2 Pre-Enrollment	Visit 1 Day 0 Baseline Sampling Visit	Visit 2, Day 30 – 365 Visit 3, Day 60 - 548 Re-Sampling Visits (for subjects who are available for re-sampling)
Pre-screening contact : Review study information sheet with interested individual		X		
Obtain informed consent		X		
Reaffirm consent			X	X
Assign screening number		X		
Perform/review history and physical exam		X	X	X
Vital signs or oral temperature (See Section 6.2)		X	X	X
Record concomitant medications		X	X	X
Vaginal pH		X	X	X
Collect blood for HIV, HBV, HCV		X		
Urine pregnancy test for females		X	X	X
Counsel subjects on refraining from using products and preparation for sampling visit		X		X
Instruct subjects on stool collection and provide stool collection kit		X		X
Inform subjects of a third sampling possibility and obtain contact information if subject agrees to be contacted				X
Specimen Collection	Body Site Specimen Collection		X	X
	Blood Specimen Collection		X	
	Obtain Stool specimen		X	X
Subsequent Research Laboratory Processes	Microbial Nucleic Acid Isolation, Quantitation, and Evaluation		X	X
	Microbial Nucleic Acid Sequence Generation and Data Deposition in Repositories		X	X
Other Procedures	Blood to Coriell Institute for human DNA isolation/storage, later lymphoblastoid cell line transformation; serum stored in clinical lab		X	

6.2 Detailed Description of Study Procedures

Screening:	Days -30 to -2

- Obtain written informed consent for screening and sampling signed by the subject. The consent process must be documented in the source documentation.
- Assign screening identification number.
- Measure height and weight for determination of Body Mass Index.
- Obtain vital signs: oral temperature, pulse, blood pressure.
- Assess general medical history including demographics.
- Review oral, skin, respiratory tract, digestive tract, and in females, gynecological history.
- Complete history-directed physical examination and direct examination of the oral cavity, skin, and vagina to verify study eligibility. Obtain vaginal pH (subject must not be menstruating). See Manual of Procedures.
- Record any concomitant medications.
- For female subjects, record all contraceptives used within the preceding 12 months, as described in the Manual of Procedures.
- Collect blood for HIV, HBV and HCV testing.
- In female subjects, perform urine pregnancy test. The test must be negative for subject to be eligible for inclusion.

- Following verification of study eligibility, with the proviso that HIV, HBV and HCV results may alter that eligibility, schedule subjects to return to the clinic no sooner than 48 hours later for body site sampling. For females, the baseline visit should be scheduled for a time when the subject is not menstruating.

- Remind subjects that the following drugs listed in the exclusion criteria must not have been taken in the past six months, and may not be taken between the screening visit and the baseline visit:

    - systemic antibiotics, antifungals, antivirals or antiparasitics (intravenous, intramuscular, or oral);
    - oral, intravenous, intramuscular, nasal or inhaled corticosteroids;
    - cytokines;
    - methotrexate or immunosuppressive cytotoxic agents;
    - large doses of commercial probiotics consumed (greater than or equal to 10^8 cfu or organisms per day) - includes tablets, capsules, lozenges, chewing gum or powders in which probiotic is a primary component. Ordinary dietary components such as fermented beverages/milks, yogurts, foods do not apply.
    - for female subjects, combination hormone vaginal ring for contraception (due to unknown duration of local hormone effects).

- Remind subjects that nasally-delivered live, attenuated cold-adapted influenza vaccines must not be administered during the 28 days prior to sampling.

- Remind subjects that topical antibiotics or topical steroids must not be used on the skin of the face, scalp, neck, arms, forearms or hands during the 7 days prior to sampling.

- Remind female subjects that vaginal and vulvar medications, including antifungals, must not be used during the 7 days prior to sampling. Subjects may continue to use permitted vaginal contraceptives until 48 hours prior to sampling.

- Counsel subjects to refrain from using any antibacterial/antiseptic products (as listed in the Manual of Procedures) for 48 hours prior to the baseline visit.

- Counsel subjects to refrain from swimming in a chlorinated pool or using a hot tub for 48 hours prior to the baseline sampling visit. Counsel female subjects to refrain from douching, vaginal sexual activity, and use of contraceptive spermicides, diaphragms, cervical caps, contraceptive sponges, suppositories, feminine sprays and genital wipes for 48 hours prior to the baseline visit.

Subjects may remain on other forms of contraception throughout study enrollment.

- Counsel subjects to not bathe or shower or brush their teeth or floss their teeth within 12 hours of the scheduled baseline appointment. Hand washing is permissible.

- Provide stool collection kit; instruct subject in the use of the kit and return of the specimen at baseline visit.

Visit 1 - Baseline Sampling:	Day 0

- Reconfirm subject's willingness to participate in core microbiome specimen collection.
- Verify inclusion and exclusion criteria.
- Collect medical history data.
- Record concomitant medications that have changed since the screening visit. Confirm that subject has not taken drugs listed in exclusion criteria during the last six months.
- Confirm that subject has not received a nasally-delivered live, attenuated cold-adapted influenza vaccine during the last 28 days.
- Confirm that subject has not used topical antibiotics or topical steroids on the skin of the face, scalp, neck, arms, forearms or hands during the previous 7 days.
- Confirm that female subject has not used vaginal and/or vulvar medications, including antifungals, during the previous 7 days. Subjects may have used permitted vaginal contraceptives until 48 hours prior to sampling.
- Confirm that female subject is not menstruating.
- Measure oral temperature to screen out subjects with febrile infectious disease.
- In females, perform urine pregnancy test. The test must be negative before the specimens are collected.
- Confirm that subject has abstained for 48 hours from use of antibacterial/antiseptic products as listed in the Manual of Procedures.
- Confirm that subject has abstained for 48 hours from swimming in a chlorinated pool or use of a hot tub; for female subjects, confirm that subject has abstained from douching, vaginal sexual activity, and use of contraceptive spermicides, diaphragms, cervical caps, contraceptive sponges, suppositories, feminine sprays and genital wipes in previous 48 hours. Subjects may remain on other forms of contraception throughout study enrollment.

- Confirm that subject has abstained for 12 hours from bathing or showering or brushing teeth or flossing teeth.
- In females, measure the vaginal pH.
- Collect blood and body site specimens (oral cavity, skin, nasal cavity, vagina) according to instructions in Manual of Procedures.
- Retrieve stool specimen brought in by the subject.

Visit 2 - Re-sampling:	Day 30 – 365

This visit will occur for subjects who are re-contacted and agree to be sampled again.

- Following the re-contact, schedule an appointment for re-sampling and provide a stool specimen collection kit in advance to the subject. Remind subject of products and activities to be avoided for 48 hours prior to sampling. For females, the re-sampling visit should be scheduled for a time when the subject is not menstruating.

- Reconfirm subject's willingness to participate in core microbiome specimen collection.

- Inform subject of the possibility of a third sampling and obtain contact information if subject agrees to be contacted for a possible third sampling.

- Collect current medical history data with particular attention to any changes since the screening medical history.

- Record any concomitant medications. Confirm that subject has not taken drugs listed in exclusion criteria during the last six months. The re-sampling may be delayed until six months after the last use of these medications, provided that the second sampling occurs within one year of the first sampling.

- Confirm that subject has not received a nasally-delivered live, attenuated cold-adapted influenza vaccine during the last 28 days. The resampling may be delayed until 28 days after the vaccine administration, provided that the second sampling occurs within one year of the first sampling.

- Record intercurrent use of systemic antibiotic, antifungal, antiviral or antiparasitic drugs (for time period between baseline sampling visit and resampling visit).

- Confirm that subject has not used topical antibiotics or topical steroids on the skin of the face, scalp, neck, arms, forearms or hands during the previous 7 days.

- Confirm that female subject has not used vaginal and/or vulvar medications, including antifungals, during the previous 7 days. Subjects may have used permitted vaginal contraceptives until 48 hours prior to sampling.

- Confirm that female subject is not menstruating.
- Measure oral temperature to screen out subjects with febrile infectious disease.
- Complete a history-directed physical examination.
- In females, perform urine pregnancy test. The test must be negative before the specimens are collected.
- Confirm that subject has abstained for 48 hours from use of antibacterial/antiseptic products as listed in the Manual of Procedures.
- Confirm that subject has abstained for 48 hours from swimming in a chlorinated pool or use of a hot tub; for female subjects, confirm that subject has abstained from douching, vaginal sexual activity, and use of contraceptive spermicides, diaphragms, cervical caps, contraceptive sponges, suppositories, feminine sprays and genital wipes in previous 48 hours. Subjects may remain on other forms of contraception throughout study enrollment.
- Confirm that subject has abstained for 12 hours from bathing or showering or brushing teeth or flossing teeth.
- In females, measure the vaginal pH.
- Collect body site specimens (oral cavity, skin, nasal cavity, vagina) according to instructions in Manual of Procedures.
- Retrieve stool specimen brought in by the subject.

Visit 3 - Re-sampling:	Day 60 – 548

This visit will occur for 100 subjects (50 from each study site) who are re-contacted and agree to be sampled a third time.

- Following the re-contact, schedule an appointment for re-sampling and provide a stool specimen collection kit in advance to the subject. Remind subject of products and activities to be avoided for 48 hours prior to sampling. For females, the re-sampling visit should be scheduled for a time when the subject is not menstruating.

- Confirm that subject has signed the informed consent form that describes the third specimen collection.

- Collect current medical history data with particular attention to any changes since the screening medical history.

- Record any concomitant medications. Confirm that subject has not taken drugs listed in exclusion criteria during the last six months.

- Confirm that subject has not received a nasally-delivered live, attenuated cold-adapted influenza vaccine during the last 28 days. The resampling may be delayed until 28 days after the vaccine administration, provided that the third sampling occurs within six months of the second sampling.

- Record intercurrent use of systemic antibiotic, antifungal, antiviral or antiparasitic drugs (for time period between second and third sampling visits).

- Confirm that subject has not used topical antibiotics or topical steroids on the skin of the face, scalp, neck, arms, forearms or hands during the previous 7 days.

- Confirm that female subject has not used vaginal and/or vulvar medications, including antifungals, during the previous 7 days. Subjects may have used permitted vaginal contraceptives until 48 hours prior to sampling.

- Confirm that female subject is not menstruating.

- Measure oral temperature to screen out subjects with febrile infectious disease.

- Complete a history-directed physical examination.

- In females, perform urine pregnancy test. The test must be negative before the specimens are collected.

- Confirm that subject has abstained for 48 hours from use of antibacterial/antiseptic products as listed in the Manual of Procedures.

- Confirm that subject has abstained for 48 hours from swimming in a chlorinated pool or use of a hot tub; for female subjects, confirm that subject has abstained from douching, vaginal sexual activity, and use of contraceptive spermicides, diaphragms, cervical caps, contraceptive sponges, suppositories, feminine sprays and genital wipes in previous 48 hours. Subjects may remain on other forms of contraception throughout study enrollment.

- Confirm that subject has abstained for 12 hours from bathing or showering or brushing teeth or flossing teeth.

- In females, measure the vaginal pH.

- Collect body site specimens (oral cavity, skin, nasal cavity, vagina) according to instructions in Manual of Procedures.

- Retrieve stool specimen brought in by the subject.

6.3 Specimen Collection, Preparation, Handling and Shipping

6.3.1 Specimen collection

Oral Cavity: Specimens will be collected from multiple locations within the mouth, saliva, soft tissue and hard tissue sites. Saliva will be collected from the floor of the mouth with a calibrated pipette. Soft tissue specimens will be collected from the tongue dorsum, hard palate, buccal mucosa, keratinized (attached) gingiva, palatine tonsils and throat using sterile Catch-All™ Sample Collection Swabs. Supragingival plaque and subgingival plaque will be collected from a minimum of four molar teeth using a Gracey curette. See the Manual of Procedures for details of specimen collection.

Skin and Nasal Cavity: Specimens will be collected from the retroauricular crease of both ears and from the antecubital fossa (inner elbow) of both arms, by rubbing with sterile Catch-All™ Sample Collection Swabs. Specimens will be collected from the anterior right and left nares, using a sterile applicator. See the Manual of Procedures for details of specimen collection.

Gastrointestinal tract: Stool from a single bowel movement will be collected by the subject. Subjects will be provided with a materials kit for collecting stool prior to the baseline visit. Subjects will bring collected stool to the baseline visit. Stool will be collected within 24 hours before the baseline visit and kept cool using the provided frozen gel packs and Styrofoam box. See the Manual of Procedures for details of specimen collection.

Vagina: Specimens will be collected from the posterior fornix of the vagina, the midpoint of the vagina, and the vaginal introitus using sterile Catch-All™ Sample Collection Swabs. See the Manual of Procedures for details of specimen collection.

Blood: Whole venous blood specimens will be collected using appropriate aseptic technique. At the baseline sampling visit, 30 mL of blood will be collected in three 10-mL tubes. Two tubes will be used to collect and store whole blood for human DNA isolation and possible later creation of cell lines. These specimens will be sent to Coriell Institute for Medical Research. One tube will be used to collect blood from which serum will be separated and stored at the clinical site. The serum may be used for future research studies related to the human microbiome (such as measurements of immune responses to organisms identified in the microbiome), and, with the subject's permission, for studies not necessarily related to the human microbiome. See the Manual of Procedures for details of specimen collection.

6.3.2 Specimen handling, preparation and nucleic acid extraction

See the Manual of Procedures for protocols for handling the clinical specimens and extracting microbial nucleic acids.

6.3.3 Specimen shipment

See the Manual of Procedures for instructions regarding transport or shipment of specimens to the genome sequencing centers and to the blood specimen repository.

7 UNANTICIPATED PROBLEMS AND SERIOUS ADVERSE EVENTS

The methods of specimen collection in this microbiome sampling study pose only minimal risk to the study subjects. As defined in 45 US Code of Federal Regulations (CFR) 46.102 (i), “Minimal risk means that the probability and magnitude of harm or discomfort anticipated in the research are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests.” The minimal physical risks associated with the sampling procedures are described in Section 2.3.1 and in the informed consent document.

The nature of the information collected from subjects may involve risk to their privacy. These risks are described in Section 2.3.1 and in the informed consent document. Appendices A and B contain tables and a flowchart that address the disposition and access issues for original specimens as well as the isolated DNA, possible cell lines, and sequence data that may be derived from the original specimens.

The results of research cannot be foreseen, so it is possible that unanticipated problems may arise in the study. In addition to unexpected adverse events, there are other types of incidents, experiences, and outcomes that occur during the conduct of human subjects research that represent unanticipated problems but are not considered adverse events. For example, some unanticipated problems involve social or economic harm instead of the physical or psychological harm associated with adverse events. The investigator or designee is responsible for the detection and documentation of unanticipated problems and Serious Adverse Events (SAE) in subjects participating in the study. Unanticipated problems and SAEs should be reported to the Clinical Data Coordinating Center (CDCC) and to the IRB as outlined in the following sections and in the Manual of Procedures.

7.1 Definition of an Unanticipated Problem

The Office for Human Research Protections (OHRP), Department of Health and Human Services (DHHS) considers unanticipated problems, in general, to include any incident, experience, or outcome that meets **all** of the following criteria:

- unexpected (in terms of nature, severity, or frequency) given (a) the research procedures that are described in the protocol-related documents, such as the IRB-

approved research protocol and informed consent document; and (b) the characteristics of the subject population being studied;

- related or possibly related to participation in the research (*possibly related* means there is a reasonable possibility that the incident, experience, or outcome may have been caused by the procedures involved in the research); and

- suggests that the research places subjects or others at a greater risk of harm (including physical, psychological, economic, or social harm) than was previously known or recognized.

An incident, experience, or outcome that meets the three criteria above generally will warrant consideration of substantive changes in the research protocol or informed consent process/document or other corrective actions in order to protect the safety, welfare, or rights of subjects or others. Examples of corrective actions or substantive changes that might need to be considered in response to an unanticipated problem include:

- changes to the research protocol initiated by the investigator prior to obtaining IRB approval to eliminate apparent immediate hazards to subjects;
- modification of inclusion or exclusion criteria to mitigate the newly identified risks;
- implementation of additional procedures for monitoring subjects;
- suspension of enrollment of new subjects;
- suspension of research procedures in currently enrolled subjects;
- modification of informed consent documents to include a description of newly recognized risks; and
- provision of additional information about newly recognized risks to previously enrolled subjects.

7.2 Reporting of Unanticipated Problems

Institutions engaged in human subjects research conducted or supported by DHHS must have written procedures for ensuring prompt reporting to the IRB, appropriate institutional officials, and any supporting department or agency head of any unanticipated problem involving risks to subjects or others (45 CFR 46.103(b)(5)). Furthermore, for research covered by an assurance approved for federalwide use by OHRP, DHHS regulations at 45 CFR 46.103(a) require that institutions promptly report any unanticipated problems to OHRP.

Incidents or events that meet the OHRP criteria for unanticipated problems require the completion of an unanticipated problem report form. OHRP recommends that investigators include the following information when reporting an adverse event, or any other incident, experience, or outcome as an unanticipated problem to the IRB:

- appropriate identifying information for the research protocol, such as the title, investigator's name, and the IRB project number;
- a detailed description of the adverse event, incident, experience, or outcome;
- an explanation of the basis for determining that the adverse event, incident, experience, or outcome represents an unanticipated problem;
- a description of any changes to the protocol or other corrective actions that have been taken or are proposed in response to the unanticipated problem.

For multicenter research protocols, if a local investigator at one institution engaged in the research independently proposes changes to the protocol or informed consent document in response to an unanticipated problem, the investigator should consult with the study sponsor or coordinating center regarding the proposed changes because changes at one site could have significant implications for the entire research study.

Forms describing unanticipated problems will be submitted to the CDCC and the IRB. The CDCC will notify the NIH study representative of the unanticipated problem. Unanticipated problems will be reported immediately to the NIH Clinical Studies Oversight Committee (CSOC). The CSOC may convene an *ad hoc* meeting for review of the unanticipated problem and consideration of corrective actions. Other supporting documentation of the problem may be requested by the CSOC and should be provided as soon as possible. A summary of the CSOC meeting and proposed actions will be provided to the study Steering Committee. The study PIs will then submit the CSOC review and any protocol changes to the IRBs.

Investigators will be responsible for reporting unanticipated problems from the time that the first subject is enrolled until one year after the final subject is enrolled.

7.3 Definition of a Serious Adverse Event

The OHRP, DHHS defines an Adverse Event (AE) as any untoward or unfavorable medical occurrence in a human subject, including any abnormal sign (for example, abnormal physical exam or laboratory finding), symptom, or disease, temporally associated with the subject's participation in the research, whether or not considered related to the subject's participation in the research (modified from the definition of adverse events in the 1996 International Conference on Harmonisation E6 Guidelines

for Good Clinical Practice). A Serious Adverse Event/Experience (SAE) is any adverse event/experience that meets any of the following criteria:

- Results in death
- Is life-threatening (places the subject at immediate risk of death from the event as it occurred)
- Requires inpatient hospitalization or prolongation of existing hospitalization
- Results in a persistent or significant disability or incapacity
- Results in congenital anomaly/birth defect
- Any other adverse event that, based upon appropriate medical judgment, may jeopardize the subject and may require medical or surgical intervention to prevent one of the other outcomes listed in this definition.

For those events meeting the definition of Serious Adverse Events, the completion of an SAE form is required.

7.4 Reporting of Serious Adverse Events

OHRP considers adverse events that are unexpected, related or possibly related to participation in research, and *serious* to be the most important subset of adverse events representing unanticipated problems, because such events always suggest that the research places subjects or others at a greater risk of physical or psychological harm than was previously known or recognized, and routinely warrant consideration of substantive changes in the research protocol or informed consent process/document or other corrective actions in order to protect the safety, welfare, or rights of subjects. Furthermore, OHRP notes that IRBs have authority to suspend or terminate approval of research that, among other things, has been associated with unexpected serious harm to subjects (45 CFR 46.113). In order for IRBs to exercise this important authority in a timely manner, they must be informed promptly of those adverse events that are unexpected, related or possibly related to participation in the research, and serious (45 CFR 46.103(b)(5)).

All serious adverse events will be:

- recorded on the appropriate serious adverse event case report form;
- followed until satisfactory resolution or until the Principal Investigator or Subinvestigator deems the event to be chronic or the patient to be stable;

- reported to the IRB and to the CDCC;
- entered in the CDCC data management system.

Any adverse event considered serious by the Principal Investigator or Sub-investigator or which meets the aforementioned criteria must be submitted on an SAE form to the CDCC. Investigators will be responsible for reporting SAEs that occur during sampling or within 48 hours of a study visit.

The study clinician will complete a Serious Adverse Event Form within the following timelines:

- All deaths and immediately life-threatening events, whether related or unrelated, will be recorded on the Serious Adverse Event Form and sent by fax within 24 hours of site awareness.
- Serious adverse events other than death and immediately life-threatening events, regardless of relationship, will be reported via fax within 72 hours of site awareness.

The CDCC will notify the NIH study representative of the serious adverse event within one business day. SAEs that are related to study participation will be reported immediately to the NIH CSOC. The CSOC may convene an *ad hoc* meeting for review of the SAE and consideration of corrective actions. Other supporting documentation of the event may be requested by the CSOC and should be provided as soon as possible. A summary of the CSOC meeting and proposed actions will be provided to the study Steering Committee. The study PIs will then submit the CSOC review and any protocol changes to the IRBs.

Serious adverse events that are considered unrelated to study participation will be provided to the CSOC in a line listing to be reviewed at regularly scheduled meetings.

8 STATISTICAL CONSIDERATIONS

8.1 Study Outcome Measures

This study aims to discover the microbial diversity found within specimens collected from different individuals, but from the same regions of the body. It will also contribute to knowledge about microbial diversity within individuals over time. The amount of microbial diversity will be measured by the identification of said microbes found in the nucleic acids isolated from the body site specimens. It should be noted, however, that it may not be possible in all cases to identify all microbes contributing to the diversity of the subject.

8.2 Sample Size Considerations

The study is exploratory in nature and the number of samples planned for it is pragmatically derived. Originally, the study aimed to enroll approximately 250 subjects. Based on preliminary analysis of data, the total enrollment goal was increased to 300 subjects, to provide more robust data on the composition of the microbiome at multiple body sites. Preliminary database searches at the collection sites found a significant number of candidate subjects available for this study.

8.3 Subject Enrollment and Follow-Up

This study will consent and enroll approximately 300 subjects. Subjects who are sampled at two visits could be enrolled in the study for up to 1 year. Approximately 100 subjects (50 at each site) who are sampled a third time could be enrolled in the study for up to 18 months. It is anticipated that subject retention will rely on the interest of the subjects in the study as well as on reimbursement for their time and effort.

8.4 Analysis Plan

Analysis of microbial sequence data will be conducted by the study chair and others as needed. These efforts will include analysis of 16S rRNA sampling data used to define the microbial “species” isolated in each specimen. Once identified, these sequences can be used to assess the diversity found both within specimens isolated from the same individual and across specimens isolated from the same body regions from multiple individuals. All sequence data collected will be made available for analysis. Spurious data will be isolated and removed as DNA sequence assessment tools allow for

identification of said sequences. Missing data are not significant because the sample size was pragmatic. As such, any sequence information gained from any of the specimens will be found to contribute to the overall understanding of the microbial diversity found in the human population. Genotyping or sequencing of human DNA isolated from the blood of subjects who provide clinical specimens may be used to examine the relationship of host genotype to the microbiota present on/in an individual.

9 ACCESS TO SOURCE DATA/DOCUMENTS

Each participating site will maintain appropriate research records for this trial, in compliance with Section 4.9 of ICH E6 GCP, and regulatory and institutional requirements for the protection of confidentiality of subjects. As part of participating in a NIH-sponsored study, each site will permit authorized representatives of the sponsor and regulatory agencies to examine (and when required by applicable law, to copy) clinical records for the purposes of quality assurance reviews, audits and evaluation of the study safety and progress.

10 ETHICS AND REGULATORY CONSIDERATIONS

10.1 Principles Governing Study Conduct

This study will be conducted according to the principles of respect for persons, beneficence, and justice as stated in the Belmont Report. This study will also embrace the principles set forth in the Declaration of Helsinki. Investigators will comply with the provisions for the protection of the rights and welfare of human research subjects set forth in the U.S. Code of Federal Regulations, Title 45 Part 46, and in compliance with determinations of all Institutional Review Boards (IRBs) overseeing the research. All institutions participating in the protocol will have in place a Federal Wide Assurance (FWA) with the DHHS Office of Human Subjects Research Protections. The assurance documents the institution's commitment to the human subjects regulations.

The study will also be conducted under Good Clinical Practice (GCP) as laid out in the International Conference on Harmonisation (ICH) E6 GCP Consolidated Guidance (ICH 1996). Investigator responsibilities are set out in Section 4 of the E6 Guideline (as published in the Federal Register May 1997). Sponsor responsibilities are set out in Section 5 of the E6 ICH Guideline (as published in the Federal Register May 1997).

10.2 Institutional Review Board (IRB)

Only those IRB members who are independent of the investigator should provide opinion on a study-related matter.

This protocol (unaltered and in its entirety) and any other documents that the IRB may need to fulfill its responsibility, including subject-recruitment procedures, and information about payment and compensation available to subjects, will be submitted to an appropriate IRB by the investigator or designate, and their written unconditional approval should be in the possession of the investigator and the sponsor before commencement of the study.

No deviations from, or changes to, the protocol should be initiated without prior written IRB approval/favorable opinion of an appropriate amendment, except when necessary to eliminate immediate hazards to the subjects or when the change(s) involves only logistical or administrative aspects of the study. Protocol administrative changes are submitted to the IRB for information only. However, written verification that a memo describing the protocol administrative change was submitted should be obtained.

The IRB must be informed by the investigator (or designate) of the following:

- all subsequent protocol amendments, informed consent changes or revisions of other documents originally submitted for review - requires IRB approval;
- all subsequent protocol administrative changes - for information;
- new information that may affect adversely the safety of subjects or the conduct of the study;
- an annual update and/or request for re-approval - Continuing Review - requires IRB approval;
- when the study has been completed.

10.3 Informed Consent

Each institution will use a separate version of the study informed consent, developed from a template provided by the NIH. All subjects screened must document informed consent by signing and dating the study consent form. Those subjects found ineligible by the screening process for enrollment will not be scheduled for the baseline specimen collection visit. Subjects who agree to be sampled a third time will sign an informed consent document that includes information about the third specimen collection. Subjects may be contacted in the future and be invited to participate in future sampling studies.

The written consent document embodies the elements of informed consent as described in the Declaration of Helsinki and in the U.S. Code of Federal Regulations, Title 45 Part 46, and will adhere to the ICH Harmonized Tripartite guidelines for Good Clinical Practice. Informed consent should be implemented before any protocol-specific procedures are carried out. Information will be presented orally and in written form.

An investigator or designate will describe the protocol to potential subjects face to face. The Subject Information and Consent Form may be read to the subjects, but, in any event, the investigator will give the subjects ample opportunity to inquire about details of the study and ask any questions before dating and signing the consent form. All illiterate individuals will have the Subject Information and Consent Form explained to them point by point by the interviewer in the presence of an impartial witness. The witness should personally sign and date the consent form.

Subjects must be informed that the study involves research, and must be given information about the purpose of the study, expected benefits, possible risks, the expected duration of the subject's participation in the study, the subject's responsibilities, the approximate number of subjects involved in the study, and the procedures associated with the research study. Subjects will be informed that they will be notified in a timely manner if information becomes available that may be relevant to

their willingness to continue participation in the study. Subjects must receive an explanation as to whether compensation and any medical treatments are available if injury occurs and, if so, what that consists of, or whether further information may be obtained. Subjects must be informed of the anticipated financial expenses, if any, to the subject for participating in the study. They must be informed whom to contact for answers to any questions relating to the research project (e.g., the investigator), relating to the research subject's rights, and relating to a research-related injury (e.g., the IRB, an ombudsman, an ethics committee, or other informed administrative body). The subjects must be informed that participation is voluntary and that they are free to withdraw from the study for any reason at any time, without penalty or loss of benefits to which the subject is otherwise entitled. Neither the investigator nor the study staff should coerce or unduly influence a subject to participate or continue to participate in the study. The extent of confidentiality of subject records must be defined, and subjects must be informed that applicable data protection legislation will be complied with. Subjects must be informed that the monitor(s), auditor(s), IRB, NIH, and the regulatory authorities will be granted direct access to the subject's original medical records for verification of clinical study procedures and/or data, without violating the confidentiality of that subject, to the extent permitted by the applicable laws and/or regulations, and that by signing a written consent form, the subject is authorizing such access. Subjects must be informed that records identifying the subject will be kept confidential, and to the extent permitted by the applicable laws and/or regulations, will not be made publicly available and, if the results of the study are published, the subject's identity will remain confidential.

The consent forms will be formatted according to Institutional requirements using a model provided by NIH. Consent forms will be in a language fully comprehensible to the prospective subjects. Informed consent shall be documented by the use of a written consent form approved by the IRB and signed and dated by the subject, and by the person who conducted the informed consent discussion. The signature confirms that the consent is based on information that has been provided to the subject and that there has been opportunity to clarify this information. Each subject's signed consent form must be kept on file by the investigator for possible inspection by the Regulatory Authorities and/or NIH. The subject should receive a copy of the signed and dated written consent form and any other written information provided to the subjects, and should receive copies of any signed and dated consent form updates and any amendments to the written information provided to subjects.

10.4 Exclusion of Children (Special Populations)

Children under 18 are excluded from this study because an adult representative population is being sought in order to establish a baseline representative catalogue of the microbiome.

10.5 Subject Confidentiality

Protected Health Information (PHI) is health information that identifies a subject. PHI is protected by federal law under the Health Insurance Portability and Accountability Act (HIPAA). To take part in this research, subjects must give the research team permission to use and disclose (share) their PHI for the study explained in the consent form for this protocol.

A separate Certificate of Confidentiality (CoC) from the US Department of Health and Human Services will be sought by each Institution recruiting, consenting and sampling subjects. This will help to further protect information that may identify subjects. The CoC prevents the investigator from being forced to disclose identifying information for use in court. The investigator may not even be forced by court subpoena. Courts that may be prevented from getting subjects' information include any federal, state, local, civil, criminal, administrative, legislative, or other court proceeding. Subjects will be informed that a CoC does not prevent them or a member of their family from voluntarily releasing information about themselves or their involvement in this research. The investigator may not withhold information if subjects give their insurer or employer permission to receive information about their participation in this research. This means that they and their family must also actively protect the subjects' privacy. The CoC does not prevent the researchers from taking steps, including reporting to authorities, to prevent serious harm to the subject or others.

The consent form documents that the research team may share a subject's information with:

- The Department of Health and Human Services (DHHS) to complete federal responsibilities for audit or evaluation of this study;
- Public health agencies to complete public health reporting requirements;
- Hospital or University representatives, to complete Hospital or University responsibilities for oversight of this study;
- Primary care physician if a medical condition that needs urgent attention is discovered;
- Appropriate authorities to the extent necessary to prevent serious harm to the

subject or others.

Subject confidentiality is strictly held in trust by the participating investigators, their staff, and the sponsor(s) and their agents. The microbial sequence data will be placed in an open access database available on the internet, but reasonable efforts will be made to remove all traces of human sequence data before these data are posted, so it will be virtually impossible for anyone to identify a subject from looking at these data. The subjects' medical data, and any data generated from the sequencing of the subjects' blood specimens, will be placed in a controlled access database available on the internet. An NIH Data Access Committee will control access to these data, to ensure that the data are made available to qualified researchers, consistent with the terms of the informed consent. No information concerning the study or the data will be released to any unauthorized third party without prior written approval of the sponsor. The study monitor or other authorized representatives of the sponsor may inspect all documents and records required to be maintained by the Investigator, including but not limited to, medical records (office, clinic or hospital) and pharmacy records for the subjects in this study. The clinical study site will permit access to such records.

10.6 Future Use of Stored Specimens

It is anticipated that the original clinical specimens collected from the oral cavity, skin, nasal cavity, and vagina will be completely consumed in the DNA extraction process. It is anticipated that the original stool specimen will be in excess of the amount needed for DNA extraction. The primary focus of this study is extraction and sequencing of microbial DNA. If the volume of clinical specimen allows, it is possible that RNA extraction may be undertaken from the same specimens for RNA sequencing and gene expression studies. Any body site specimens remaining after completion of nucleic acid extraction and sequencing will be destroyed. Specimens will be stored temporarily to enable sufficient quality assessment (amount, purity) of extracted nucleic acid prior to sequencing. Nucleic acid may be re-extracted from the same clinical specimen(s) if sufficient material remains intact after the initial extraction. Any residual microbial DNA or RNA remaining after sequencing will be stored at the clinical research sites. Coded human DNA and (possibly) lymphoblastoid cell lines from the blood specimens will be maintained after the study is completed. Coded serum specimens will also be maintained after study completion.

The human DNA, the lymphoblastoid cell lines, the serum, and the microbial nucleic acids obtained under this project may be used for future approved studies related to the human microbiome. The repositories will require all researchers who apply to use the human DNA, the lymphoblastoid cell lines and the microbial nucleic acids to submit a written description of the research, which will be reviewed to make sure that the

researcher requesting the specimens plans to use them in a way that is consistent with the subject consent (i.e., related to human microbiome research). It is anticipated that researchers from universities, hospitals, non-profit groups, companies, and government research laboratories in the United States and in other countries will study the specimens and the information from them.

Coded serum specimens will be stored in repositories at the clinical research sites and may be distributed to other researchers for future studies related to the human microbiome, and with the permission of the subject, to other researchers for future studies not necessarily related to the human microbiome. The clinical site repositories will require all researchers who apply to use the serum specimens to submit a written description of the research intent and obtain approval for use of the serum.

If a subject chooses to withdraw consent for future use of stored specimens, specimens will be destroyed and every attempt will be made to exclude data from that subject in future releases of DNA sequence data. However, if data have been released to researchers prior to the subject's withdrawal of consent, data cannot be retracted.

11 INTERNAL AND EXTERNAL QUALITY CONTROL AND QUALITY ASSURANCE

The clinical collection site maintains quality control via standard operating procedures as defined by the protocols found in this document, protocols in the Manual of Procedures and in Institutional Human Research Subject SOPs. Each institution will establish internal quality control systems to ensure that from the start of screening, subject enrollment and specimen collection, the protocols are implemented consistently and with a high degree of accuracy.

The NIH will implement an external quality assurance program to assess the integrity of the site internal quality control and assurance programs and to audit the informed consent process, specimen tracking and storage, and assurance of confidentiality of documents and data. The findings of the internal programs and the NIH audits will be conveyed to the sites for investigator action.

12 DATA HANDLING AND RECORD KEEPING

The EMMES Corporation will serve as the Clinical Data Coordinating Center (CDCC) for this study, and will be responsible for data management, quality review, and reporting of the study data. Forms for use as source documents will be derived from the eCRFs and will be provided by the CDCC to the sites to record and maintain data for each subject screened and/or enrolled in the study. Data reported in the eCRF will be verified for consistency with the source documents and any discrepancies will be documented. Guidance for investigators on completing source documents and making corrections to the source documents and eCRFs can be found in the Manual of Procedures.

12.1 Confidentiality

Precautions are in place to ensure accurate record keeping procedures and to maintain subject confidentiality. Following clinical specimens collection, all of the specimens and information will be processed according to the procedures in the Manual of Procedures.

12.2 Data Management Responsibilities

The responsibilities of data management will be handled by the clinical site, the study CDCC, and, for the blood specimens, the Coriell Institute. The specific responsibilities of the clinical site include accurate collection and storage of all information pertaining to the specimens obtained for this project. The clinical site will code each specimen and maintain this information in a secure location, as described in the Manual of Procedures.

All source documents and laboratory reports will be reviewed by the clinical team and data entry staff, who will ensure that they are accurate and complete. Serious adverse events must be assessed for causality and reviewed by the site Principal Investigator or designee. Data collection will be the responsibility of the study staff at the site under the supervision of the site Principal Investigator. During the study, the Investigator will maintain complete and accurate documentation for the study.

The laboratories isolating nucleic acids will have standard lab practices for record keeping and take steps to make sure specimen integrity is maintained. The laboratories will also maintain and distribute specimen specific metrics used for specimen evaluation.

Responsibilities for the CDCC include cumulative record keeping of all information associated with the project including coded subject information collected at the time of consent, specimen specific metrics, DNA sequence information and the resulting analysis information. At the end of the study, a copy of all datasets will be provided to the NIH. The CDCC will send a copy of all datasets to the NIH electronically.

Only coded information will be transferred between the entities involved in this study. This coded information will be shared using a secure file transfer protocol. Only the study investigators at the clinical sites will have access to the study identification code list.

12.3 Data Capture Methods

Data capture methods used for this study include the following: paper copies that will be kept centrally and processed on an ongoing basis, electronic spreadsheets that will be distributed and processed on an ongoing basis, and databases that will be kept centrally as well as distributed and updated on an ongoing basis.

Clinical data (including SAEs and concomitant medications) will be entered into a 21 CFR Part 11-compliant internet data entry system (IDES) provided by The EMMES Corporation. The data system includes password protection and internal quality checks, such as automatic range checks, to identify data that appear inconsistent, incomplete, or inaccurate. Clinical data will be entered directly from the source documents.

12.4 Types of Data

The following types of coded data will be collected:

- Subject's health history;
- Specimen specific metrics such as DNA quantity and quality;
- DNA sequence from each specimen.

As described in the Manual of Procedures, subject ID codes will be assigned to subjects sequentially as they are screened. The Subject ID will be entered on all specimens and source documents and will be used to identify all subject data records.

12.5 Timing/Reports

A Steering Committee (see Manual of Procedures) will hold regular discussions via conference calls on the overall strategy and progress of the project. The information discussed on these calls will be used to generate reports used to document the overall

progress of the study. The Steering Committee will review the progress of specimen collection, DNA isolation, and analysis of the sequence data.

Sequencing metrics will also be reported to NHGRI on a quarterly basis. Information such as the number of subjects, sequence pass rates and total bases of sequencing information will be included in these documents. Information from these reports will also be provided to the Steering Committee for their review.

The collection sites are responsible for preparing clinical study reports and providing them to the appropriate regulatory authorities, according to the applicable regulatory requirements.

12.6 Study Records Retention

The Clinical Principal Investigator at each site must maintain adequate and accurate records as specified in Essential Documents for the Conduct of a Clinical Trial (Section 8 of the ICH E6 Guideline) to enable the conduct of the study to be fully documented and the study data to be subsequently verified. These documents should be classified into two separate categories: (1) Investigator's study file and (2) Subject clinical source documents. Essential documents should be retained at least seven years after the grant is officially closed or until the NIH authorizes transfer or destruction of study records. No study records will be destroyed without prior authorization from the NIH.

12.7 Repeat Specimen Collection for Nucleic Acid Extraction

Repeat specimen collection will be performed if additional nucleic acid extraction is necessary and the clinical specimen(s) is exhausted. Additional nucleic acid extraction might be necessary because of compromised clinical specimens (inappropriate storage at home, in the lab, etc.), the presence of amplification/sequencing inhibitors in particular specimens, insufficient nucleic acid yields or DNA/RNA of poor quality. Technical errors during nucleic acid extraction may also result in yields of low amount or quality.

If DNA amount and purity are questionable, as detected by quality assessment procedures in the clinical laboratories, and the source clinical specimen is exhausted, the clinical staff at the site will be notified as soon as possible. In this situation, one of two options may apply: 1) the same subject could be resampled, or 2) new subjects could be screened and enrolled in the study. If subjects are contacted for repeat body site sampling, the specimen collection will be targeted so that only the body sites that yielded inadequate nucleic acid preparations will be re-sampled.

Additional individuals may need to be recruited if enrolled subjects that are re-contacted are no longer available or refuse to provide additional specimens. Subjects may be unavailable for different reasons including health reasons (exclusion criteria) or geographic relocation. In these cases, new individuals will be screened and considered for enrollment in the study. These subjects would be newly consented for complete body sampling including the sites of interest that yielded inadequate nucleic acid preparations. The study Steering Committee, in consultation with the local IRB, will make final decisions on whether the total numbers of subjects and specimens need to be reconsidered due to specific issues with DNA collection or funding constraints.

12.8 Protocol Deviations

A protocol deviation is any noncompliance with the clinical trial protocol, Good Clinical Practice (GCP), or Manual of Procedures requirements. The noncompliance may be on the part of the subject, the investigator, or the study site staff. As a result of deviations, corrective actions are to be developed by the site and implemented promptly.

These practices are consistent with Good Clinical Practice:

4.5 Compliance with Protocol, sections 4.5.1, 4.5.2, and 4.5.3

5.1 Quality Assurance and Quality Control, section 5.1.1

5.20 Noncompliance, sections 5.20.1, and 5.20.2.

It is the responsibility of the site to use continuous vigilance to identify and report deviations within 5 working days of identification of the protocol deviation, or within 5 working days of the scheduled protocol-required activity. All deviations must be promptly reported to NIH, via the EMMES Corporation's IDES.

All deviations from the protocol must be addressed in study source documents. A completed copy of the NIH Protocol Deviation (PD) Form must be maintained in the Regulatory File. Protocol deviations must be submitted to the local IRB/IEC per their guidelines. The site PI/study staff is responsible for knowing and adhering to their IRB requirements.

13 ADMINISTRATIVE MATTERS

To comply with Good Clinical Practice (GCP), important administrative obligations relating to investigator responsibilities, monitoring, archiving data, audits, confidentiality and publications must be fulfilled. See the Manual of Procedures for details.

14 LITERATURE REFERENCES

Eichinger, N. 2007. European funding targets big biology. *Nature* **445:** 8-9.

Forney, L.J., J.A. Foster, and W. Ledger. 2006. The vaginal flora of healthy women is not always dominated by Lactobacillus species. *J Infect Dis* **194:** 1468-1469; author reply 1469-1470.

Gill, S.R., M. Pop, R.T. Deboy, P.B. Eckburg, P.J. Turnbaugh, B.S. Samuel, J.I. Gordon, D.A. Relman, C.M. Fraser-Liggett, and K.E. Nelson. 2006. Metagenomic analysis of the human distal gut microbiome. *Science* **312:** 1355-1359.

Gordon, J.I., R.E. Ley, R. Wilson, E. Mardis, J. Xu, C.M. Fraser, and D.A. Relman. 2005. Extending Our View of Self: the Human Gut Microbiome Initiative (HGMI). *www.genome.gov/Pages/Research/Sequencing/SeqProposals/HGMISeq.pdf*.

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

Luna, R.A., L.R. Fasciano, S.C. Jones, B.L. Boyanton, T.T. Ton, and J. Versalovic. 2007. DNA Pyrosequencing-Based Bacterial Pathogen Identification in a Pediatric Hospital Setting. *J Clin Microbiol* **(in press)**.

Medini, D., C. Donati, H. Tettelin, V. Masignani, and R. Rappuoli. 2005. The microbial pan-genome. *Curr Opin Genet Dev* **15:** 589-594.

NIH Office of Portfolio Analysis and Strategic Initiatives. 2007. Human Microbiome Project. *http://nihroadmap.nih.gov/hmp*.

NIH Roadmap Human Microbiome Project. 2007. Human Microbiome Project Initiatives. *http://nihroadmap.nih.gov/hmp/initiatives.asp*.

Paustian, T. 2006. The normal flora of humans. *www.bact.wisc.edu/Microtextbook*.

Tettelin, H., V. Masignani, M.J. Cieslewicz, C. Donati, D. Medini, N.L. Ward, S.V. Angiuoli, J. Crabtree, A.L. Jones, A.S. Durkin et al. 2005. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". *Proc Natl Acad Sci U S A* **102:** 13950-13955.

Thies, F.L., W. Konig, and B. Konig. 2007. Rapid characterization of the normal and disturbed vaginal microbiota by application of 16S rRNA gene terminal RFLP fingerprinting. *J Med Microbiol* **56:** 755-761.

Tierno, P.M. 2001. *The Secret Life of Germs*. Pocket Books, New York.

Turnbaugh, P.J., R.E. Ley, M.A. Mahowald, V. Magrini, E.R. Mardis, and J.I. Gordon. 2006. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444:** 1027-1131.

Weinstock, G., R. Gibbs, R. Wilson, J. Gordon, S. Clifton, B. Birren, and C. Nusbaum. 2007. Pilot Project to Expand the Number of Sequences of Culturable Microbes from the Human Body.

www.genome.gov/Pages/Research/Sequencing/SeqProposals/HMPP_Proposal.pdf.

Weinstock, G., R. Gibbs, R. Wilson, J. Gordon, S. Clifton, B. Birren, D. Ward, and C. Nusbaum. 2007. Pilot Experiments for Metagenomic Sample Sequencing.

APPENDIX A: DISPOSITION OF BLOOD AND BODY SITE SPECIMENS, CELL LINES, DNA, AND SEQUENCE DATA

Disposition of Screening Blood Specimen

Specimen Type	**Blood** 10 mL (2 tsp) at Screening Visit			
Purpose	**Screening for HIV, HBV and HCV to determine study eligibility**			
Disposition Issue	**Original Specimen**	**Cell Line**	**DNA (genomic)**	**Data from specimens** **DNA sequence information (genomic)**
Storage – short term	Specimen will be stored in clinical lab and will be destroyed after screening tests are completed.	N/A	N/A	N/A
Banking	N/A	N/A	N/A	N/A
Who has access?	Study personnel and lab personnel will collect and test specimen; subject will receive test results, but they will not be entered in subject's medical record.	N/A	N/A	N/A
Revoked authorization?	Specimen will be destroyed.	N/A	N/A	N/A

Disposition of Blood Specimens, Cell Lines, Human DNA, Human Sequence Data

Specimen Type	**Blood** 30 mL (6 tsp) at Visit 1 Baseline Sampling
Purpose	**Genome sequencing and development of cell lines** **Storage of serum for future use in human microbiome-related studies (or in unrelated studies, with subject's consent)**

Disposition Issue	**Original Specimen**	**Cell Line**	**DNA (genomic)**	**Data from specimens DNA sequence information (genomic)**
Storage – short term	Study personnel will code whole blood specimens and ship them to Coriell Institute for temporary storage, DNA extraction, and cell line development. Serum will be coded and stored at sites.	N/A	N/A	N/A
Banking	Coded whole blood specimens will be stored at Coriell Institute until they are processed and lymphocytes are cryo-preserved. Serum will be stored at sites for possible future research.	Coded lymphoblastoid cell lines may be developed and banked indefinitely at Coriell Institute repository.	Extracted DNA will be stored indefinitely as coded samples at Coriell Institute repository.	Genetic sequences will be stored indefinitely.
Who has access?	Serum specimens will be stored in controlled access repositories for future research; specimens will not be available to the subject or the subject's physicians for other testing.	Any cell lines created will be stored in Coriell controlled access repository, and may be distributed to researchers after Coriell, in consultation with NIH (NHGRI), approves research intent; cell lines will not be available to the subject or the subject's physicians.	Coded DNA samples will be stored in a controlled access repository, and may be distributed to researchers after Coriell, in consultation with NIH (NHGRI), approves research intent; DNA samples will not be available to the subject or the subject's physicians for other testing.	Human sequence data will be available to qualified researchers only through a controlled-access internet database; neither the subject nor the subject's physicians will receive sequence information.
Revoked authorization?	Remaining coded blood and serum specimens will be destroyed.	Coded cell lines in repository will be destroyed; attempt will be made to retrieve and destroy any cell lines already distributed to researchers.	Coded DNA samples remaining in repository will be destroyed; attempt will be made to retrieve and destroy any DNA samples already distributed to researchers.	Sequence data will be deleted from future database versions; however, data previously released on the internet database cannot be retracted.

Disposition of Body Site Specimens, Microbial DNA*, Microbial Sequence Data

Specimen Type	**Oral cavity swabs** **Skin specimens** **Nasal swabs** **Stool specimens** **Vaginal swabs**
Purpose	**Metagenome sequencing**

Disposition Issue	**Original Specimen**	**DNA (metagenomic)**	**Data from specimens** **DNA sequence information (metagenomic)**
Storage – short term	Specimens will be stored and processed for DNA extraction at the site clinical lab/tissue procurement bank.	N/A	N/A
Banking	N/A (Any specimen remaining after nucleic acid extraction and sequencing will be destroyed.)	Coded samples of DNA for metagenome sequencing will be stored indefinitely in controlled access repositories at the clinical sites.	Genetic sequences will be stored indefinitely.
Who has access?	Clinical lab/tissue procurement bank personnel will have access to specimen for processing.	Coded DNA samples will be stored in a controlled access repository; they will be provided to sequencing centers for metagenome sequencing and may be distributed to other researchers after the clinical center (Baylor College of Medicine or Washington University) approves research intent; DNA samples will not be available to the subject or the subject's physicians.	Microbial sequence data will be available through an open-access internet database.
Revoked authorization?	If consent is withdrawn prior to DNA extraction, specimen will be destroyed. No specimen will remain after nucleic acid extraction and sequencing.	Coded DNA samples remaining in repository will be destroyed; attempt will be made to retrieve and destroy any DNA samples already distributed to researchers.	Sequence data will be deleted from future database versions; however, data previously released on the internet database cannot be retracted.

***and possibly microbial RNA, depending on specimen availability**

APPENDIX B: HMP SPECIMEN FLOWCHART

Collection, Processing, Storage of Specimens and Sequence Data
Disposition of stored/banked specimen if subject withdraws consent