The human microbiome and cancer: New opportunities for population studies

Emily Vogtmann, PhD, MPH
Research Fellow
Metabolic Epidemiology Branch
Division of Cancer Epidemiology & Genetics
National Cancer Institute

November 18, 2016
Helicobacter pylori and gastric cancer

- Primary etiologic agent for gastric cancer
- First bacterium determined to cause cancer
- Unlikely to discover another single bacterium to cause cancer
The human microbiome and cancer

- The human microbiome, mainly from oral and fecal samples, has been found to be associated with a number of cancers, including:
  - Colorectal cancer
  - Esophageal cancer
  - Hepatobiliary cancers
  - Lung cancer
  - Pancreatic cancer
- Replication of findings has been an issue
- Nearly all studies utilized a case-control design
Priorities for epidemiologic studies

- Preservation of the microbial signature or “biomarker” in the field over days in suboptimal storage conditions

- Optimization of collection method for multiple technologies

- Quality control standards to evaluate reproducibility

- Collection of new microbiome samples in order to study incident diseases
Outline

- Collection methods for microbiome studies
- Quality control (QC) samples
- Developing new population studies
Outline

- Collection methods for microbiome studies
- Quality control (QC) samples
- Developing new population studies
Fecal collection for 16S rRNA gene studies

- Mayo I (Sinha R, et al, 2016 *CEBP*)
  - 20 individuals
  - Only fecal samples
  - Frozen immediately, ambient temperature for 1 or 4 days

- Mayo II (Vogtmann E, et al, Accepted, *AJE*)
  - 53 individuals
  - Fecal and oral samples
  - Frozen immediately or ambient temperature for 4 days
Fecal collection methods

No Solution → RNA later → 95% Ethanol
Evaluation criteria

- Technical reproducibility: Compare replicates

- Stability: Compare Day 4 to Day 0

- Accuracy: Compare to “gold standard”
  - Fecal sample: No additive, Day 0
Microbiome characteristics

- Relative abundance of 3 phyla
  - Actinobacteria
  - Bacteroidetes
  - Firmicutes
- Alpha diversity
  - Observed OTUs
  - Shannon Index
- First principal coordinate of beta diversity matrices
  - Unweighted, generalized, and weighted UniFrac
  - Bray-Curtis
Statistical analysis

- Intraclass correlation coefficient (ICC)
  - Compares the variability within a subject to the variability across subjects

- Spearman correlation coefficient
  - Non-parametric test
  - Compares the rank order of individuals
Overall variability explained

Vogtmann E, et al, Accepted, AJE
Technical reproducibility

Day 0

ICC

Actinobacteria, Bacteroidetes, Firmicutes, Observed OTUs, Shannon index, UniFrac PC1, GUniFrac PC1, WUniFrac PC1, BC PC1

No Solution, FIT Tube, FOBT Card, RNALater, 95 ethanol

Vogtmann E, et al, Accepted, AJE
Accuracy (ICC)

Vogtmann E, et al, Accepted, AJE
Accuracy (Spearman)

![Accuracy Graph]

Vogtmann E, et al, Accepted, AJE
Summary

- Interindividual variability greatly outweighed differences by collection method or freezing timepoint
- All methods appeared to be reproducible, stable, and relatively accurate
- Future studies could use these methods for 16S rRNA gene analyses
  - However, comparisons should be made within a collection method
Oral collection methods

OMNIgene DISCOVER

Picture 1

Vogtmann E, et al, In preparation
Fecal metabolomics collection study

Technical reproducibility

Stability

Accuracy

Summary

- Both 95% ethanol and FOBT cards were relatively reproducible, stable, and accurate compared to the “gold standard”

- FIT tubes performed less well for untargeted metabolomics

- The immediately frozen, no additive fecal sample had high detectability and the highest estimates of technical reproducibility

- New studies can use 95% ethanol and FOBT cards for both 16S rRNA gene sequencing and untargeted metabolomics
Outline

- Collection methods for microbiome studies
- Quality control (QC) samples
- Developing new population studies
Quality control samples

- QC sample type 1: Artificial (mock) community
  - Oral artificial community: 22 species
  - Gut artificial community: 20 species
- QC sample type 2: Robogut
DNA extraction pilot

Sample types:
- Oral samples
- Artificial communities

Vogtmann E, et al, Unpublished
DNA extraction pilot

Vogtmann E, et al, Unpublished
Summary

- QC samples are important for epidemiologic studies
  - Batch effects
  - Data pooling
- Need for more complex, known QC sample in proper matrix
Outline

- Collection methods for microbiome studies
- Quality control (QC) samples
- Developing new population studies
Considerations for new microbiome studies

- Study design
  - Hypothesis
  - Validated cases
  - Appropriate controls
  - Statistical power
- Specimen collection
  - Body site, sub-site, or surrogate
  - Quantity
  - Contamination
  - Stabilization

Goedert JJ (2013) *Eur J Clin Invest*
Considerations for new microbiome studies

- Specimen handling and analysis
  - Processing, lysis, and extraction
  - Primer selection
  - Sequencing depth
  - Data processing and assignment of taxa

- Confounders
  - Antibiotics
  - Other medications
  - Smoking
  - Diet

Goedert JJ (2013) *Eur J Clin Invest*
Prospective cohort study
Collection of microbiome samples in prospective cohort studies

- Baseline collection
  - Oral/fecal biospecimen collection
  - Comprehensive questionnaire
- Follow-up(s)
  - Collection of additional oral/fecal biospecimens
  - Additional questionnaires
- Identification of endpoints using cancer registry and National Death Index
- Conduct nested case-cohort study when sufficient cases accrued
Overall conclusions

- Many collection methods are available for fecal samples for microbiome analyses
  - More work is needed for other ‘omic technologies
  - More work is needed for other samples types (e.g., tumor tissue)

- Important need to develop QC samples for inclusion in all microbiome studies

- Collection of new samples for prospective microbiome studies are essential to understand the impact of the microbiome on health and disease
Acknowledgements

- DCEG Metabolic Epidemiology Branch
  - Christian Abnet
  - James Goedert
  - Erikka Loftfield
  - Gwen Murphy
  - Rashmi Sinha

- DCEG Biostatistics Branch
  - Jianxin Shi
  - Xing Hua
  - Yunhu Wan

- DCEG Cancer Genomics Research Laboratory
  - Belynda Hicks
  - Amy Hutchinson
  - Kristie Jones
  - Elisa Lee
  - Shalabh Suman
  - Bin Zhu

- University of Maryland
  - Owen White

- UCSD
  - Rob Knight
  - Amnon Amir

- University of Guelph
  - Emma Allen-Vercoe

- Mayo Clinic
  - Nicholas Chia
  - Jun Chen

- Harvard University
  - Curtis Huttenhower