Microbiome

Michael Snyder

February 27, 2018
What is the microbiome?

- The collection of microbiomes that exist in a given area, tissue or organism. It refers to all microbes including bacteria, fungi and viruses.
Hot topic

From Ami Bhatt
The Human Microbiome

• The human microbiome is the collection of microbes that live on and in us

  10-30 trillion human cells
  1.6-10X microbial cells
  3% of human body mass
  >10 million genes (compared to 20K for humans)

Enormous biochemical diversity

What do the microbes do??
The Role of the Microbiome

• It breaks down our food into constituents we absorb.

• It synthesizes essential vitamins. E.g. Vitamin B12, Biotin, Folic Acid
Body ecology

Different niches
- Aerobic, anaerobic
- Dry, moist
- Oily, nutrient-rich/poor
- Exposed, protected from environment

Different Kingdoms
- Bacteria
- Eukaryota (e.g. fungi)
- Archaea
- Viruses

Different Phyla/Classes
- Typically 5-10 major types
- 100s – 1000s of species
- Macro-ecology: Plants, Animals, Fungi, Protists
Where did our microbiome come from?

- Microbes have been on earth for 3 billion years

We evolved in the midst of microbes
- We evolved organs to house our microbes
- Each body site has its own microbial community
- Adaptive immunity allowed richer microbiome?
When does a person’s microbiome form?

- At birth a baby’s gut microbiome is similar to the mother’s vaginal microbiome.
- It forms a lot during the first few years as baby ingest food.
- Microbiome from breast fed microbiome has similarities to its Mother’s breast/milk.
- Non breast fed babies have a different microbiome.

Dominguez-Bello et al. (2010) PNAS 107:119
Sampling of the microbiota 5-60 minutes after birth

Vaginal (Baby like Vaginal) Caesarian (baby looks like skin)

Dominguez-Bo et al. (2010) PNAS107:11971 x
Approaches to metagenomic study of the microbiome

- Microbial community is very diverse
- Bacteria, viruses, eukaryotes
- 100s-1000s of taxa
- A few are >80% of total
- MANY minor taxa
- 99% of microbes cannot be cultured (under aerobic conditions)!
Targeted sequencing

16S rDNA = bar code identifier of species

- 16S rDNA sequences/sample
  - Average 5000 sequences
  - Must have >1000 sequences
  - Sample 100s of taxa

Describe communities and genus abundance in many samples – “Average” community and variations

Cost < $50/sample

Primers to variable parts of 16S rDNA

Liu et al. Nucleic Acid. Res. 2007 35. e120
Shotgun sequencing (metagenomics)

- Sample every gene in the community by high throughput DNA sequencing

Analyze for:
- Bacterial community structure, Viruses, Fungi
- Metabolic pathways, other genes of interest (antibiotic resistance; virulence)

Slide courtesy of Ami Bhatt
Growth of Archae/Bacteria genome sequences

1 genome/day added to GenBank

Bacterial sequencing dominates

Nikos Kyrpides, Genomes Online Database
Lots of diversity of individual level: People Have Their Own “Personal Microbiome”

Major subgroups of the stool biome type

Histograms of genera in each sample

Firmicutes/Ruminococcus

Prevotella

Each row a different sample

Bacteroides

Yanjiao Zhou & George Weinstock
Weight gain/loss Samples Cluster by Subject, not Perturbation

Legend:

- Bacteroidaceae
- Ruminococcaceae
- Lachnospiraceae
- Unclassified
- Porphyromonadaceae
- Rikenellaceae
- Enterobacteriaceae
- Verrucomicrobiaceae
- Sutterellaceae
- Erysipelotrichaceae
- Peptostreptococcaceae
- Coriobacteriaceae
- Veillonellaceae
- Prevotellaceae
- Streptococcaceae
- Acidaminococcaceae
- Clostridiaceae\n
Family Composition
Diversity at the organismal level but wide array of shared microbial genes comprising a “core microbiome” at the gene level

The microbiome differs with different organ sites
Other cool facts

The skin microbiome is associated with our attractiveness to mosquitos!

HA = Highly Attractive
PA = Poorly attractive

Does microbiome affect attraction between humans?

---

Niels O. Verhulst PLoS One 2011
The Role of the Microbiome in Health and Disease
Diet and Health: Two Very Different Metabolic Scenarios

High Fiber Diet

Small intestine

Colon

Circulation

SCFA

SCFA

Diet and Health: Two Very Different Metabolic Scenarios

High Simple Carbohydrate Diet

(Sonnenburg and Sonnenburg, Cell Metab., 2014)
Some examples: Obesity and the microbiome

- Our food affects our microbiome
- High fat vs low fat diet.
Less gut bacterial diversity in obese vs lean people

Obesity and the microbiome

- Differences between obese mouse and lean mouse

Ley et al PNAS 2005
Balance Between Genome, Immune System and Microbiome
TLR5 (toll like-receptor deficient) mice fight bacteria differently from normal mice.
Their transmissible microbiome makes them eat more

Vijay-Kumar et al Science 2010
What about humans?
Shift in microbiome during weight loss

Ley et al Host Cell & Microbe 2006
Study in humanized mice
Diabetes and the microbiome

• Diabetics have a different microbiome than nondiabetics in humans and mouse models.

• Inject bacteria from healthy mouse into mouse prior to inducing diabetes. Reduces diabetes formation.

Wang et al Cell 2012
The Microbiome Partially Predicts Glucose Response (Segal et al. 2015)
Ecological diseases from our microbiome

• Not infectious diseases: causative agents are part of the normal flora
  - e.g. *Candida albicans*
  - Disease results from disturbance to the environment, altering microbial balance

• Antibiotic treatment - *Clostridium difficile* no longer held in check by excess beneficial organisms
  - Reduced protection against deleterious organisms

• Dental caries - *Streptococcus mutans* overgrowth due to high carbohydrate utilization
  - Increased amounts of deleterious substances, eg lactic acid

• Toxic Shock Syndrome - *Staphylococcus aureus* emergence due to abnormally dry environment
  - Change in the environment

Many more - to be discovered
Disease metagenomics projects

- Illness of unknown origin
- Necrotizing enterocolitis
- IBD/Crohn’s disease
- Urethritis
- Nasal microbiome after vaccination and flu infection
- Acne
- Antibiotic effects on physiology
- Cystic fibrosis
- Periodontitis
- Lung microbiome and HIV
- Kawasaki disease
- Bacterial vaginitis
- Obesity
- Diabetes
- Sepsis
- Trachoma and sequelae
- Autism
Myocardial infarction (heart attack)

High levels of TMAO (trimethylamine-N-oxide) are associated with MI

- Carnitine from meat is converted to TMAO using gut microbiome
- Antibiotics suppress TMAO levels in humans
- Germ free mice, no TMAO

- CUT genes responsible

**Possible solutions:**
- 1) Change diet to reduce relevant bacteria
- 2) Produce inhibitors of the CUT gene product.
- 3) Genetic engineering

http://guardianlv.com/2013/10/flu-vaccine-may-cut-heart-attack-risk/
Autism Spectrum Disorders (ASD)

• Altered microbiome in ASD patients
  Increased: Clostridia and Desulfovibrio in stools

• Increased SCFAs in ASD patients: acetic, propionic, and butyric acid

• Adding PPA induces repetitive behaviour in mice

• Probiotics improved ASD in mice and maybe humans?

Ceymi Doenyas Neuroscience 374, Pages 271-286
Inflammatory bowel disease (IBD)

- Different microbiome in Crohn’s disease and IBD
- In some cases fecal transplants seemed to help
  - (In case you were wondering –
  - Administered through ingestion or the rectum)
Clostridium difficile associated diarrhea

Kulpers & Surawicz, Lancet 2008
Capsulized FMT

• >70% effective with first administration
• ~92-94% effective with second administration
• Transportable, storable, quality controlled
• No procedure required
Infections: Organ transplantation

Microbial infections often appear after organ transplantation

Example: Hematopoietic cell transplantation with umbilical cord cells at Harvard 11 people infected after transplant

e.g. 30 yr old woman with ALL. Pursued cord transplant.

Developed coiltis

WBC 135K/uL, 53% blasts → Acute Leukemia (ALL)

Bhatt et al, NEJM 2012
Organ transplantation

Extract DNA from fixed tissue → Shotgun sequencing

Many reads did not map to known bacterial genomes!

Bhatt et al, NEJM 2012
Organ transplantation

Assemble unmapped reads from two patients into new sequence: *Bradyrhizobium enterica*

Slide courtesy of Ami Bhatt
Classification of the CCS-associated bacterium

Phylogenetic analysis using the draft genome to classify the organism

Slide courtesy of Ami Bhatt
Conclusions

1) The microbiome is vast and essential

2) It is associated with many, and possibly all, human diseases.

3) It is not clear what is cause and what is effect.

4) It may provide novel avenues for treating human disease.
Metabolomics
Metabolome: Definition

The collection of small molecules (typically less than 1500 Daltons) that are in a cell, tissue or organism.

There are 1000s of metabolites in an organisms—exact number is not known.
Examples of famous metabolites:

**Endogenous**
- ATP (energy source)
- cAMP (activates Protein kinase A)
- Diacylglycerol (PKC pathway)
- Cholesterol (affect membrane properties and signalling pathways)

**Exogenous**
- Ethanol
- Resveratrol (Anti-aging)
- Morphine
- Nicotine
Why study metabolomes?

Metabolomes are closer to phenotypes.
Why study metabolomes?

Metabolomes are closer to phenotypes.

In total numbers, there are many more metabolites than other types of molecules.

<table>
<thead>
<tr>
<th>Environment</th>
<th>Genes</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>mRNA</td>
<td>16-50</td>
<td>~10^6</td>
</tr>
<tr>
<td>Protein</td>
<td></td>
<td>5 x 10^7</td>
</tr>
<tr>
<td>metabolite</td>
<td></td>
<td>3 x 10^{12}</td>
</tr>
<tr>
<td>Phenotype</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Metabolites come from:

Us (the things we make): e.g.
Nucleosides and nucleotides (DNA, RNA building blocks)
Most membrane lipids

Our microbiome
Short chain fatty acids (SCFA), Biotin, Folate

Foods we eat
½ of our amino acids (e.g. Phenylalanine, Tryptophan, Lysine)
Most vitamins (A, D, E, K, B)
Omega 3 fatty acids
Metals e.g. Zn\(^{2+}\), Fe\(^{2+}\), Mn\(^{2+}\)
Caffeine
We also get metabolite *precursors* from our food

Beta carotene converts to Vitamin A retinol

Ergosterol: sunlight converts to Vitamin D2 (ergocalciferol)

Only source of vitamin D for vegetarians
Diverse functions of metabolites:

• Perform the biochemical reactions of a cell
  Energy, e.g. ATP

• Form building blocks of macromolecules (e.g. Proteins, Nucleic Acids, Glycans are formed from amino acids, nucleotide)

• Serve as cofactors of enzymes (e.g. NAD, Heme)

• Regulate enzyme activity
  Acetyl-CoA → Acetylation
  ATP → Phosphorylation
Consequently, they are involved in diverse cellular processes:

Cell signalling (e.g. ligand for receptors)
  cAMP: beta adrenergic receptor
  Prostaglandin: Prostaglandin receptors
  Caffeine: adenosine receptors
  Sildenafil (Viagra): cGMP-specific phosphodiesterase type 5

Cell proliferation and cell cycle progression
  Warburg effect (uncoupled glycolysis and respiration)
  Serine, essential to prostate cancer (non-essential to normal cells)
There are a wide range of metabolic diseases

e.g. Genetic diseases such as
- Phenylketourea
- Diabetes
- Obesity
- Lipidemias
- Cancer (Warburg effect)
The metabolome is less well studied than other omes (transcriptome, proteome, genome)
The human metabolome is large

We do not know how big it is.

Few thousand to 10s of thousands
The metabolome is chemically diverse which makes it hard to study (Unlike other macromolecules like nucleic acids and proteins)

Hydrophobic  
- e.g. fatty acids, lipids, cholesterol

Hydrophilic Positive  
- e.g. amino acids, amines

Hydrophilic negative  
- e.g. sugars, organic acids

Hydrophobicity (LogP)  
Molecular Weight (Da)
How to study metabolomes

1) Profiling: Mass Spectrometry
   Liquid chromatography (LC-MS): Untargeted and targeted
   Gas Phase Mass Spectrometry (GC-MS): Targeted

2) Interactions:
   a) Arrays of metabolites positioned in an addressable format
   b) Pull down experiments
Metabolite profiling by LC-Mass spectrometry

Untargeted LC-MS
Similar to proteomics

Extract metabolites:
50-80% methanol
(extracts hydrophobic and hydrophilic)


LC separation

MS acquisition

[M+H]+

MS spectrum of L-Methionine

Pathway analysis & reconstruction

Identification & quantitation of metabolites

Features selection
Using sophisticated LC, many thousands of MS peaks are detected in blood and urine.

Peaks detected in urine and blood using:

- Reverse phase LC (for hydrophobic molecules)
- HILIC LC (for hydrophilic molecules)

Run each sample in positive and negative modes to analyze negative and positive ions, respectively.

Only $1/5^{th}$ to $1/10^{th}$ of peaks are identified.

HILIC separation

MS acquisition
Using sophisticated LC, many thousands of MS peaks are detected in blood and urine.

**URINE – ESI (+) MS**
- 8462 metabolic features
- HILIC 29%
- RPLC 39%
- 11,332 unique features in both ESI modes

**URINE – ESI (-) MS**
- 8854 metabolic features
- HILIC 32%
- RPLC 31%

**PLASMA – ESI (+) MS**
- 5188 metabolic features
- HILIC 40%
- RPLC 43%

**PLASMA – ESI (-) MS**
- 4739 metabolic features
- HILIC 60%
- RPLC 24%

**PLASMA**
- 8,287 unique features in both ESI modes
Protein profiling by LC-Mass spectrometry: Targeted metabolomics

• Similar to proteomics
• Selectively targets analytes of interest.
• Focuses the instrument on selected targets, greatly enhancing signal-to-noise.
• Can be coupled to stable isotope dilution.
• Enables precise quantification.
Metabolite profiling by LC-Mass spectrometry: Targeted metabolomics

**LC-MS**
- Metabolite mixture
  - Liquid chromatography
  - Select precursor ions m/z
  - Detect fragment ion m/z

**LC-MS/MS**
- Fragment precursor ions
  - Detect fragment ion m/z

**Liquid Chromatography**
- Metabolite mixture
  - LC
  - Metabolites selection
  - Fragmentation (for MS/MS)
  - Fragment selection

**MRM signal**
- Intensity vs. Time
Profiling by LC-Mass spectrometry: Targeted metabolomics


Rat liver homogenate
Applications of metabolomics

1) Discover biomarkers

1) Link metabolic abnormalities to genetic mutations

2) Understand regulatory networks and pathways
Discover biomarkers: Cardiovascular research

1983 patients 63 metabolites

Identify short chain acylcarnitines
Ser, Gly, FFA
Discover biomarkers: Type 2 Diabetes

Profile:
Obese BMI 37 kg/m²
vs Lean 23 kg/m²

Identify branched-chain amino acids: leucine/isoleu and valine, glutamate/glutamine (glx) and C3 and C5 acylcarnitines, and the aromatic amino acids phenylalanine and tyrosine

Insulin Resistance vs Insulin Sensitive (~60 people)
- 1791 named metabolites
- Welch t test on IS and IR (median of all baselines)
- 108 significant metabolites with FDR < 0.2 (qvalue)

Amino acids, acylcarnitines, indolelactic acid, tetrahydrocortisol glucuronide
Discover biomarkers (example)

3) Cancer

**Key Cancer Types**

**Sample of Key of Metabolite Differences**

(Healthy Controls/ Benign Disease vs Malignancy)

**Breast**
- Phosphocholine
- Choline
- Taurine
- Glucose
- 5-hydroxymethyl-2-deoxyuridine
- 8-hydroxy-2-deoxyuridine
- Choline Metabolism
- Amino Acid Metabolism
- Energy Metabolism
- Modified Nucleosides: Oxidative DNA Damage

**Ovarian**
- 3-hydroxy-butyrate
- Lactic acid
- Glycerolphosphate alpha
- Phosphoric acid
- Inositol-2-phosphate
- Uric acid
- Glutamic acid
- Proline
- Glycine
- Free Fatty Acid

**Colon**
- Beta-alanine
- Taurine
- Methionine
- Uric acid
- Oleic Acid
- Free Fatty Acid
- Inositol Phosphate Metabolism
- Choline containing compounds
- Succinate
- Isocitrate
- Gut Flora Metabolism

Sawyer, J. Surg. Oncol. 2010; 21831: 5
Discover biomarkers (example)

3) Cancer

Serine depletion reduces tumor growth and promotes mouse survival

Maddocks et al. Nature. 2013
Recent Study Linking Metabolites and Human Disease: Combine Metabolomics Profiling with WES

Metabolomic profiling of 80 people: 575 diverse metabolites, 72 pathways
## Whole Exome Sequencing

<table>
<thead>
<tr>
<th>Sample/Person</th>
<th>Pathway</th>
<th>Genes</th>
<th>Variants</th>
<th>Rare &amp; damaging</th>
</tr>
</thead>
<tbody>
<tr>
<td>3905</td>
<td>Glycolysis/gluconeogenesis</td>
<td>89</td>
<td>32</td>
<td>ALDOB p.I177L</td>
</tr>
<tr>
<td>3923</td>
<td>Purine metabolism</td>
<td>21</td>
<td>9</td>
<td>XDH p.R1296W</td>
</tr>
<tr>
<td>3890</td>
<td>Fatty acid beta oxidation</td>
<td>17</td>
<td>4</td>
<td>SLC22A5 p.V488I</td>
</tr>
<tr>
<td>3952</td>
<td>Bile acid</td>
<td>79</td>
<td>33</td>
<td>None</td>
</tr>
<tr>
<td>3926</td>
<td>BCAA metabolism, fatty acid, glycolysis</td>
<td>60</td>
<td>22</td>
<td>None</td>
</tr>
<tr>
<td>3923</td>
<td>BCAA metabolism, fatty acid, glycolysis</td>
<td>60</td>
<td>21</td>
<td>MAPK8IP1 D386E, MC4R p.I1251L</td>
</tr>
<tr>
<td>3891</td>
<td>BCAA metabolism, fatty acid</td>
<td>60</td>
<td>26</td>
<td>None</td>
</tr>
<tr>
<td>3902</td>
<td>BCAA metabolism, fatty acid, glycolysis</td>
<td>60</td>
<td>7</td>
<td>THADA p.E1741K, CLPS p.R55H</td>
</tr>
<tr>
<td>3917</td>
<td>Bile acid</td>
<td>79</td>
<td>33</td>
<td>TTC37 p.L1505S</td>
</tr>
<tr>
<td>3937</td>
<td>BCAA metabolism, fatty acid, glycolysis</td>
<td>60</td>
<td>26</td>
<td>LDLR p.P526H, FN3K p.H146R, PASK</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td><strong>p.P1249L</strong></td>
</tr>
</tbody>
</table>
Bile acid circulation and dot plots showing data distribution in the cohort (n = 80) for the four primary bile acids.

High Bile acid patients

Patient 3917: TTC37 p.LI505S Mutation: Affect circulation in liver
Patient with Carnitine Deficiency: Mutations in organic cation/carnitine transporter gene (SLC22A5 p.V488I)

Lining Guo et al. PNAS 2015;112:E4901-E4910

Lining Guo et al. PNAS 2015;112:E4901-E4910
Application #3: Understanding regulatory pathways: Example

Affinity purification of tagged proteins and analysis of bound metabolites by mass spectrometry

Protein

Mock

Lysis

IgG beads

Wash

LC-MS

Metabolite Extraction

Data Analysis

Identify Bound Metabolites

Xiyen Li et al., 2010 Cell
Yeast protein kinase often bind metabolites

- **103/122** Protein Kinases tested
- **21** bind **10** metabolites
- **48** novel interactions
- **15** bound **ergosterol**

Li et al., 2010 Cell
Kin4 binds ergosterol

LC

Untagged Solvent

MS

Untagged Solvent
Summary of affinity pulldowns

>20% of proteins bind metabolites

What do these interactions do?
Ergosterol is Required for Ypk1 Kinase Activity
Ergosterol is Required for Ssk22 Protein Levels

(a) Western blot analysis of Ssk22 and Ypk1 proteins in WT, erg4Δ, and erg4Δ + ergosterol conditions. 

(b) Time course of Ssk22 protein levels in WT and erg4Δ strains treated with ergosterol.
Kinase Interactions Are Altered by Ergosterol

WT  erg4Δ  erg4Δ  +  ergosterol
Take-home messages

1. The metabolites universe is large and comes from many sources

2. Its chemical diversity makes it harder to study
   Mass spectrometry: Targeted and untargeted

3. Wide applications:
   a) Biomarker for human diseases
   b) Identification of roles in biology
   c) Discover metabolic abnormalities and link to candidate genetic differences.