“...whoever wishes to pursue the science of medicine must first investigate the seasons of the year and what occurs in them.”

Hippocrates, 4th Century B.C.

APHL Annual Meeting
Rhode Island Convention Center
June 12, 2017

Plenary:
Advanced Molecular Detection in Environmental Matrices

Rita R. Colwell, Ph.D., D.Sc.
Distinguished University Professor
University of Maryland, College Park
and
Johns Hopkins University Bloomberg School of Public Health
## Water-related diseases

<table>
<thead>
<tr>
<th>Disease</th>
<th>Cases per year</th>
<th>Deaths per year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amoebiasis</td>
<td>48,000,000</td>
<td>110,000</td>
</tr>
<tr>
<td>Arsenic</td>
<td>28-35m exposed to drinking water with elevated levels</td>
<td></td>
</tr>
<tr>
<td>Diarrhoeal disease, Including cholera</td>
<td>1.5 billion</td>
<td>1,800,000</td>
</tr>
<tr>
<td>Dracunuliasis (guinea worm)</td>
<td>&gt; 5000</td>
<td>-</td>
</tr>
<tr>
<td>Fluorosis</td>
<td>26 million (China)</td>
<td>-</td>
</tr>
<tr>
<td>Giardiasis</td>
<td>500,000</td>
<td>Low</td>
</tr>
<tr>
<td>Hepatitis A</td>
<td>1,500,000</td>
<td>-</td>
</tr>
<tr>
<td>Intestinal helminths</td>
<td>133,000,000</td>
<td>9400</td>
</tr>
<tr>
<td>Malaria</td>
<td>396,000,000</td>
<td>1,300,000</td>
</tr>
<tr>
<td>Schistosomiasis</td>
<td>160,000,000</td>
<td>&gt; 10,000</td>
</tr>
<tr>
<td>Trachoma</td>
<td>500,000,000</td>
<td>-</td>
</tr>
<tr>
<td>Typhoid</td>
<td>500,000</td>
<td>25,000</td>
</tr>
</tbody>
</table>
Background on Cholera: A Global Disease

- Acute water-related diarrheal disease
- Seventh pandemic started in 1960s
- Occurs in more than 50 countries affecting approximately 7 million people
- Bengal Delta is known as “native homeland” of cholera outbreaks
- Since cholera bacteria
  - exist naturally in aquatic habitats
  - evidence of new biotypes emerging, *it is highly unlikely that cholera will be eradicated but clearly can be controlled by provision of safe drinking water.*
Model for the Transmission of Vibrio Cholerae from the Environment to Humans
Work begins in Bangladesh in 1975
Chronology of cholera and satellites

1996: Colwell's coastal vibrio hypotheses

2001: Lobitz's chlorophyll-SST

Several attempts to link satellite data with cholera

Classification of coastal and Inland cholera
Cholera and SST in the Indian Ocean

Six-month SST lead: $R^2 = 0.72$

Lobitz et al., 2000, PNAS Vol. 97, No. 4 pp. 1438-1443
Could we have predicted the Haiti Cholera outbreak?

- Recent cholera outbreak in Haiti indicated the disease remains a global threat.

- Framework for developing cholera prediction models in cholera endemic (ER) and non-endemic regions (NER)

- The sharp contrast in mortality rates between ER and NER exists not because we do not know how to treat cholera patients, but because of a persistent “knowledge barrier” between ER and NER.

- We propose a pragmatic and adaptive framework which hypothesizes that convergence of three enabling situations - Inception, Environmental Conditions, and Transmission - are necessary for a cholera outbreak to become an epidemic.

Air temperature in Haiti in 2010 and average of air temperature data for last 50 years.

Monthly rainfall in Haiti in 2010 compared with historical rainfall data.

Vibrio cholerae
Sequenced and published in 2000

Chromosome I (2,961,149 bp, 2,742 ORFs)

Chromosome II (1,072,315 bp, 1,093 ORFs)

Missing ORFs in V. cholerae strains (Reference: N16961; cutoff = 70% DNA similarity)

The Haitian *V. cholerae* O1 strains clustered with other 7th pandemic *V. cholerae* strains in a single monophyletic clade.
10 Haitian strains (red) form a cluster cloud, distinct and yet, distant, from CP genomes (concurrent epidemic isolates form different parts of the world) (blue) and others (green).

Interestingly, one reference strain CP 1038 (from Zambia) genome falls into the Haitian cluster.
Microbiome based Diagnostics

Microbiome Analysis for Rapid Detection of Enteric Pathogens and Characterization of the Intestinal Microbiome in Health and Disease
Microbes → Extracted DNA → DNA Sequencing

GENIUS®: Probabilistic Matching

GENEBOOK® Library

Raw Sequence Reads

Applications:
- Pathogen detection for health and wellness;
- food safety and probiotics
- biothreat and public health surveillance;
- forensic and scientific investigation
CosmosID provides analysis for shotgun metagenomics

Microbial ID using NGS Data


GenBook®: Expert-curated genome databases

Highest performance bioinformatics algorithms

Graphical User Interface

Patented Methodologies

COSMOSID®
CosmosID Value Proposition

CosmosID: Next Generation Sequencing Based Solution for Microbial Identification and Pathogen Characterization

- Incredibly fast output – in minutes (post sequencing)
- Expert-curated proprietary genome databases (GenBook®)
- High Sensitivity & High Specificity
- Sub-species and Strain level identification, including abundance
- Polymicrobial identification from a single test
- No A priori assumptions necessary
## Credentials / On-going Collaborations

<table>
<thead>
<tr>
<th>Client/Entity</th>
<th>Engagement Overview</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDA</td>
<td>• 5 Year contract as in the in-house metagenomics solution after an in depth two year bake off</td>
</tr>
<tr>
<td>NIST &amp; CosmosID Metagenomics MVP Challenge v1.0</td>
<td></td>
</tr>
</tbody>
</table>
| illumina               | • BaseSpace Import Integration  
                          • #1 all time attended webinar through the end of 2016                                                                                     |
| QIAGEN                  | • Partnering as CLC’s Metagenomics Solution                                                                                                      |
| ThermoFisher Scientific | • Partnering as Thermofisher Cloud’s Metagenomics Solution                                                                                      |
|                          |                                                                                                                                                  |
|                          | • Nominated to the White House “Cancer Moon Shot” by Vice President Joe Biden                                                                     |
| Johnson & Johnson       | • Working with Big and Small Pharma companies on discovery R&D, clinical trials and quality control                                               |
Diarrheal Disease Study:
Infectious Disease in the Genomics Era
Diarrheal Disease Study

National Institute of Cholera and Enteric Disease
Kolkata, India

www.niced.org.ind
### Enteric Pathogens Monitored By NICED
(Using non-metagenomic methods)

<table>
<thead>
<tr>
<th>Bacteria</th>
<th>Parasites</th>
<th>Viruses</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Vibrio cholerae</em></td>
<td><em>Giardia lamblia</em></td>
<td>Rotavirus</td>
</tr>
<tr>
<td><em>Vibrio parahaemolyticus</em></td>
<td><em>Cryptosporidium parvum</em></td>
<td>Adenovirus</td>
</tr>
<tr>
<td><em>Vibrio fluvialis</em></td>
<td><em>Entamoeba histolytica</em></td>
<td>Norovirus</td>
</tr>
<tr>
<td>Aeromonas spp.</td>
<td><em>Blastocystis hominis</em></td>
<td>Sapovirus</td>
</tr>
<tr>
<td>Campylobacter jejuni</td>
<td></td>
<td>Astrovirus</td>
</tr>
<tr>
<td>Campylobacter coli</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shigella</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salmonella</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Escherichia coli</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total # of NICED samples: 74</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indian Healthy Control (HC): 20</td>
</tr>
<tr>
<td>Sick with Unknown Etiology (UE): 28</td>
</tr>
<tr>
<td>Sick with Known Etiology (KE): 26</td>
</tr>
<tr>
<td>Healthy Human Microbiome Project (HMP): 20</td>
</tr>
</tbody>
</table>

In collaboration with the National Institute of Cholera and Enteric Disease (NICED), Calcutta, India
Microbiomes of Diarrheal Subjects Compared to Healthy Subjects

**Known Etiology**

**Unknwn Etiology**

**Healthy Individuals**

**Healthy Controls**

γ-Proteobacteria – Firmicutes – Bacteroidetes
Many Pathogens Readily Identified From Diseased Patients

Family Level

Healthy Control

Human Microbiome Project

Diseased, Known Etiology

Diseased, Unknown Etiology
Individuals with AMR genes present in their microbiome

Genes which match at > 50% coverage
HMP samples had no genes present which matched at this level of coverage
Microbial communities of healthy volunteers suggest that the microbiome of healthy humans of Indian descent is markedly different than those of Western European descent.

Indian population may tolerate low number of pathogenic microorganisms that may indicate a “disease state” for Western European descent

Multiple pathogens can readily be identified from disease patients

Metadata revealed that patients who exhibited profound watery diarrhea contained in their microbiome pathogens primarily of the *Escherichia coli* complex, namely pathogenic *E. coli* and *Shigella* species.

The microbial community of the Indian population encodes higher rates of antibiotic resistance genes when compared to healthy HMP samples

Functional analysis of the Indian microbiome indicates predominance of carbohydrate metabolism genes
Infectious Disease in the Genomics Era

Necrotizing fasciitis
Cross-talk among flesh-eating *Aeromonas hydrophila* strains in mixed infection leading to necrotizing fasciitis

Duraisamy Ponnusamy, Elena V. Kozlova, Jian Sha, Tatiana E. Erova, Sasha R. Azar, Eric C. FITTS, Michelle L. Kirtley, Bethany L. Tiner, Jourdan A. Andersson, Christopher J. Grim, Richard P. ISOM, Nur A. Hasan, Rita R. Colwell, and Ashok K. Chopra

*Department of Microbiology and Immunology, University of Texas Medical Branch, Galveston, TX 77555; Center for Food Safety and Applied Nutrition, Office of Applied Research and Safety Assessment, Food and Drug Administration, Laurel, MD 20708; CosmosID Inc., Rockville, MD 20850; Center for Bioinformatics and Computational Biology, University of Maryland Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742; and Bloomberg School of Public Health, The Johns Hopkins University, Baltimore, MD 21205*

Neighbor-joining tree using 2,514 conserved full length predicted proteins
Relative distribution of four *Aeromonas hydrophila* strains NF1, 2, 3, and 4 into different metagenomic datasets derived from muscle, spleen and liver samples.
Strain interaction in mixed infection

**NF1**
- **Elimination by host defense mechanism**
- Dissemination to peripheral organs
- NO ExoA, thus Intact local tissue barriers

**NF2**
- **Elimination by MΦ phagocytosis**
- Local tissue barrier damage

**Antagonism of NF1 on virulence of NF2**
- Direct and host mediated elimination

**Synergism of NF2 ExoA on NF1 dissemination**
- Dissemination to peripheral organs
• 15 *Klebsiella pneumoniae* carbapenemase (*bla*KPC)-*Aeromonas* sp. and Enterobacteriaceae clinical isolates were collected between 2/2012-5/2013.

• These were paired with 32 sink trap biofilm samples collected from the same hospital in 3/2014.

• In addition, the 15 isolates and 9 of the sink trap biofilms were sequenced.

• Metagenomic WGS data was analyzed by CosmosID.

• To determine if the patient isolates were present in any of the biofilm samples, their genomes were assembled and added to the ~31,000 genomes already in the CosmosID bacterial database.

• The biofilm WGS metagenomes were then analyzed using this supplemented database.

• Both the sink trap biofilm samples and the patient isolates were also profiled using the CosmosID antimicrobial resistance and virulence gene databases.

Collaboration with Amy J. Mathers, UVA
• An edge connects an isolate (in red) to a sink trap biofilm (in blue), if the isolate is present in the biofilm.

• An edge connects 2 isolates if they were both found in the same sink trap sample.

• The size of the isolate vertices represents the average abundance in the biofilm samples.
Except for 3 genes, all genes found in common between the isolates and the biofilms are plasmid-associated.

There are 2 distinct clusters of biofilms and isolates (circled).
The Microbiome as Therapy
Pre and Post-Fecal Microbiota Transplant (FMT) Microbial Shift

Strain associated with disease

Pre-FMT

Post-FMT
Protection of Microbiome from Antibiotic Induced Change

Prevented overgrowth of organisms known for secondary infection

Development of SYN-004, an oral beta-lactamase treatment to protect the gut microbiome from antibiotic-mediated damage and prevent *Clostridium difficile* infection
Orange County Water District Study
The influent, secondary treated municipal wastewater of the AWPF treatment train is purified by a three-step process: microfiltration, reverse osmosis, and ultraviolet (UV) light with hydrogen peroxide. Initially the wastewater is screened at 4mm Sodium hypochlorite as disinfectant is added prior to microfiltration. Hydrogen peroxide (H$_2$O$_2$) is added before UV treatment. The decarbonators remove CO$_2$ and raise the pH; addition of lime further stabilizes the purified water.
Relative abundance and diversity of parasite DNA in MF-biofilm and Q1-water. The approximate relative abundance heat map was simplified, using the GENIUS bioinformatics algorithm and curated databases. The 99 relative abundance corresponds to sequences classified as *Paramecium biaurelia* strain v14, *Thalassiosira*, and *Acanthamobea polyphaga* based on observed frequency of DNA sequences identified. Parasite sequences were not found in the RO-biofilms.
Comparison of virus and bacteriophage DNA sequences

Virus and bacteriophage DNA sequences comparison demonstrate the presence of bacteriophages and virus DNA in the membrane filter (MF)-biofilm and in the influent water, Q1. Note, the absence of bacteriophages and DNA viruses in the reverse osmosis (RO)-biofilm. Presence and absences of sequences (partial or complete) related to bacteriophage and viruses in MF-biofilm were compared to the Q1 water.
A Simple, Sustainable Method for Reducing Cholera
A Simple Solution for Cholera Prevention: Sari Filtration
Cases of Cholera Per 1000 Population

Test Group

Control

Sari

Nylon

Full Study
“When one tugs at a single thing in nature, he finds it hitched to the rest of the universe.”

John Muir
(1838-1914)
Collaborators and Colleagues

Anwar Huq, Professor
University of Maryland, College Park, MD

Antarpreet Jutla, Assistant Professor,
West Virginia University
Morgantown, WV

Dr. Nur Hasan
Vice-President, Research and Development
CosmosID, Inc.
College Park, MD

Dr. Seon Young Choi,
Bioinformatic Scientist,
CosmosID Inc.
College Park, MD
Genomics Team

**NICED:**
G. Balakrish Nair
Thandavarayan Ramamurthy

**University of Maryland:**
Shah M. Rashed
Seon Young Choi
Anwar Huq

**CosmosID:**
Nur A. Hasan
Poorani Subramanian
Kelly Moffat
Huai Li
Brian Fanelli
Manoj Dadlani

**Orange County Water District**
Menu Leddy

**Consultant**
Joseph Cotruvo

Amy J. Mathers, UVA
Ashok Chopra, University of Texas Medical Branch
Chris Grim, FDA
Safe water is a global challenge