DVD’s Diverse Disease Portfolio

- **CNS** (polio*/picornaviruses, herpesviruses, rubella)
- **Respiratory** (adenovirus, picornavirus, respiratory syncytial virus, parainfluenza virus, SARS, MERS, human metapneumovirus, mumps virus*, VZV, HSV, CMV)
- **Rash/Exanthems** (measles*, mumps*, rubella*, VZV* & other herpes, picorna, parvovirus B19)
- **Gastroenteritis** (rotavirus*, norovirus, astrovirus, adenovirus, picornavirus)

*Vaccine-preventable disease
# DVD Vaccine-Preventable Diseases

### Current
- Polio (eradication)
- Measles (elimination)
- Mumps
- Rubella/CRS (elimination)
- Varicella-Zoster
- Rotavirus

### On the Horizon?
- Respiratory Syncytial virus
- Norovirus
- Enterovirus 71 (Asia)
- Cytomegalovirus

- Polio (eradication)
- Measles (elimination)
- Mumps
- Rubella/CRS (elimination)
- Varicella-Zoster
- Rotavirus

- Respiratory Syncytial virus
- Norovirus
- Enterovirus 71 (Asia)
- Cytomegalovirus
DVD Virus Genomic Diversity

4,500 – 235,000 nucleotides

- **ssRNA(+)**: *Picornaviridae* (poliovirus, enterovirus, rhinovirus, parechovirus); *Coronaviridae* (SARS, MERS, others); *Caliciviridae* (norovirus); *Astroviridae* (astrovirus)

- **ssRNA(-)**: *Paramyxoviridae* (measles virus, mumps virus, RSV, parainfluenza viruses, metapneumovirus)

- **dsRNA**: *Reoviridae* (rotavirus) [segmented]

- **ssDNA**: *Parvoviridae* (parvovirus B19)

- **dsDNA**: *Herpesviridae* (VZV, cytomegalovirus, others); *Adenoviridae* (adenovirus)
# Relevant Clinical Specimens

<table>
<thead>
<tr>
<th>Culture</th>
<th>Clinical – Nasal Swab</th>
<th>Clinical – Stool</th>
<th>Tissue, Blood, Environmental</th>
</tr>
</thead>
</table>
| • Poliovirus  
• Rubella  
• Measles | • Rhinovirus, EV-D68  
• Rubella  
• Measles  
• Coronavirus  
• Adenovirus  
• RSV  
• Parainfluenza | • Poliovirus  
• Calicivirus  
• Picornavirus  
• Rotavirus | • Picornavirus  
• Poliovirus  
• Herpesviruses  
• Other pathogens |

**Challenge:**
- Easy
  - Take time to culture
  - Many viruses not readily cultured
- Hard
  - Low titer sample can be difficult to get enough coverage
  - Other DNAs and RNAs, inhibitors in the samples, lowering coverage of target
  - Lots of host background
Goals of DVD AMD Project

- “Right-size” NGS methods for small viral genomes
- Single wet-lab workflow for most/all DVD viruses
- Simple, yet powerful bioinformatics pipeline, usable by bench scientists
- Enhancement of VPD surveillance
Viral NGS Laboratory Pipeline

Key features

- Size selection by filtration (0.45 µm) (remove bacteria)
- Nuclease to eliminate host nucleic acid
- Random amplification
Time to Result

Miseq Sample Preparation

- **Rapid Response**
  - 1 sample
  - 13 hours total prep time
  - 2.45 hours
  - 3 hours
  - 3 hours
  - 1.45 hours
  - 2.5 hours
  - 24 hours

- **Standard**
  - 16 samples
  - 18 hours total prep time
  - 4 hours
  - 3 hours
  - 3.45 hours
  - 3.15 hours
  - 4 hours
  - 48 hours

- **High-Throughput**
  - 48-96 samples
  - 19 hours total prep time
  - 5 hours
  - 3 hours
  - 3.45 hours
  - 3.15 hours
  - 4 hours
  - 48 hours

Total: 64 hours

Ion Torrent Preparation (evaluation)

- **Emergency Response**
  - 1-4 samples
  - 11 hours total prep time
  - 2.45 hours
  - 3.45 hours
  - ~5 hours

Total: 14 hours

Slide credit: SRA International / Anna Montmayeur
Features of Bioinformatics Pipeline

- Web-based: Simple GUI, no software to install
- Automated: Monitors MiSeq and PGM output folders
- Modular: Open-source components can be upgraded, modified, exchanged
- Quality control: FastQC for each run
- Reference and de novo assembly
- Automated BLAST to curated database or GenBank
- Detailed reporting
Viral NGS Bioinformatics Pipeline

Fast Mode – Rapid Detection of DVD viruses

- Raw reads (fastq)
- Rapid Viral Surveillance
- Rapid Alignment
- Surveillance report

Comprehensive Mode – All Viral Pathogens

- Raw reads (fastq)
- De novo genome assembly
- QC sample
- Host removal
- Pre-Process Sample
- Contig alignment to ref genomes
- Contig assembly
- ID region of interest
- Extract locus
- ID proteins, indels, rearrangements & other features
- Contig based Typing
- Novel pathogen discovery
- nt to protein DB alignment

Client Specific Pipelines

Virome report

CSRA International / Eddie Ramos
Advantages to the NGS Approach

- Rapid pathogen detection/identification
  - Novel pathogens
- Better resolution for molecular epidemiologic studies
  - Identification of source of infection
  - Monitoring transmission pathways
- Monitor viral evolution (e.g. stability of epitopes and targets for diagnostic tests)
- Can generate whole genomes sequences more rapidly at reduced cost
  - Sanger sequencing has fixed cost per genome size (e.g. measles virus, 16kb, is approximately $300/genome)
  - With multiplexing, NGS methods offer substantial cost savings (1 sample=$1000/genome, 12=$144/genome, 24=$100/genome, 48=$80/genome)
Example of Improved Resolution for Molecular Epidemiology: Measles Virus

Standard window: 450nt

Whole genome: 15,894nt

Bankamp, unpublished
Alternative Method: MERS-CoV

48 samples
48 primer sets

48 barcoded libraries of 48 amplicons each

2304 amplicon sequences
Challenges: Rotavirus

- Difficult to separate dsRNA strands
- DNase treatment, LiCl purification
- Ligate “loop oligo”
- Denature with methyl mercury hydroxide
- Synthesize cDNA, remove RNA

Potgieter et al., 2009, J Gen Virol
Challenges: Measles Genome

Poor coverage in the 1000 nt non-coding region between the M and F genes

Bankamp, unpublished
NGS Efficiency vs CT Value

<table>
<thead>
<tr>
<th>Norovirus Outbreak Samples</th>
<th>% viral reads mapping to Norovirus</th>
<th>Detection of core genome</th>
<th>Detection of partial genome</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>23.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>13.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>9.0%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>8.5%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6.7%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>11.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>9.1%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>5.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>2.9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.8%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>0.4%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>0.8%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>0.1%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>0.0%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Genome Sequences, FY2015

- Poliovirus: 200
- Picornavirus: 158
- Norovirus: 29
- Measles virus: 25
- Rubella virus: 30
- Rotavirus: 67
- Coronavirus: 37
- Herpesvirus: 1
Plans for 2016-2017

- Further implement and document quality practices for AMD/NGS methods
  - SOPs
  - Laboratory protocols, instrument validation
  - Work with partners in state PHLs on methods development/validation

- Finalize Viral VPD NGS Pipeline and develop an outward-facing Viral VPD Portal

- Pilot transfer of DVD AMD standardized laboratory and bioinformatics workflows to state PHLs
  - Plan for a week-long hand-on training for State partners

- India as a Global Health Security Project

- WHO LabNet partners in 2017 (WHO NGS working group)
2016 CDC Next-Generation Sequencing of Viral Pathogens Training for State Laboratories

Dates: Monday, July 11th through Friday, July 15th
AMD Project 50 Team

- **DVD Core**
  - Terry Ng, Christina Castro, Laura Magaña, Rachel Marine, Anna Montmayeur (CSRA)

- **CSRA Bioinformatics**
  - Greg Doho, Eddie Ramos, Roman Tatusov, Yang Xu

- **Virus-specific Teams**
  - Respiratory: Dean Erdman, Xiaoyan Lu, Teresa Peret, Kapil Chandora
  - Norovirus: Jan Vinjé, Kshama Aswath, Nikki Collins, Marta Diez-Valcarce, Nicole Gregoricus
  - Rotavirus: Mike Bowen, Mathew Esona, Sunanado Roy, Kunchala Rungsrisuriyachai, Jennifer Hull, Sung-sil Moon
  - Pathogen Discovery: Sue Tong, Clint Paden, Krista Queen, Ying Tao
  - Measles: Bettina Bankamp
  - Rubella: Joe Icenogle, Min-hsin Chen
  - Herpes: Scott Schmid, Jennifer Folster
  - Picorna: Allan Nix, Ever Vega
  - Polio: Cara Burns, Qi Chen, Jane Iber, Alex Schmidt

PIs: Paul Rota, Steve Oberste
DVD AMD Laboratory Infrastructure

- AMD-dedicated staff
  - Leverage SMEs drawn from DVD lab branches
- “DVD Mini-Core Lab,” bldg. 17
  - 2 MiSeqs, 1 Ion PGM
  - Ion Chef, Zephyr, Tape Station, RainDrop dPCR
  - Dedicated pre- and post-PCR areas
- 2 MiSeqs and associated equipment in DVD labs in bldg. 18
Laboratory Method Optimization

- Pre-treatment
  - Filtration to remove bacteria
  - Nuclease treatment to enrich for encapsidated genomes
  - Nucleic acid extraction
  - Capture

- Library prep
  - Evaluate different library prep kits
  - Evaluate automated library prep
  - Different levels of multiplexing

- Illumina MiSeq vs Ion PGM

- Adapting Access.Array to other viruses
Application: Norovirus Outbreak

RdRP typing

VP1 typing

Aswath et al. in prep.