Real-time Metagenomic Analysis
(improvements in genome sequencing techniques, have they made open-ended searches for new pathogens possible?)

Christian Happi, PhD
Director, World Bank-African Center of Excellence for Genomics of Infectious Diseases (ACEGID)

Redeemer’s University, Ede, Nigeria
Metagenomics: Unbiased NGS

• It is simply running all nucleic acids in a sample, which may contain mixed populations of microorganisms, and assigning these to their reference genomes to understand which microbes are present and in what proportions.

• The ability to sequence and identify nucleic acids from multiple different taxa for metagenomic analysis makes this a powerful new platform that can simultaneously identify genetic material from entirely different kingdoms of organisms.

• The possible clinical applications are tremendous, including diagnosis of infectious diseases, outbreak tracking, infection control surveillance, and mutation and pathogen discovery, among many others.
THE 2014-2016 EBOLA VIRUS DISEASE OUTBREAK

April 2014, RUN, Nigeria
Rapid sequencing of Ebola Virus-Summer 2014: Data made public immediately.

- First large-scale genome sequence-based analysis of the circulating Ebola viral population
- June 2014: 99 genomes publicly available
- Another 150 genomes released between March 2014 and August 2015

• New deep Sequencing Methods for Lassa and Ebola viruses developed and published immediately for the benefit of the community (Matranga CB et al. Genome Biol. 15 (2014))
The data were generated super quickly!

<table>
<thead>
<tr>
<th>2 days</th>
<th>5 days</th>
<th>1 day</th>
<th>2 days</th>
</tr>
</thead>
<tbody>
<tr>
<td>extraction of host</td>
<td>library preparation</td>
<td>full-length sequence</td>
<td>assembly pathogen detection</td>
</tr>
</tbody>
</table>

Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak

Stephen K. Gire et al.  
Science 345, 1369 (2014);  
DOI: 10.1126/science.1259657

10 days from sample to Genbank
Next Generation Sequencing of Lassa Virus Genomes

- ~400 patient samples
- 22 Rodents

- Unbiased sequencing - no specific amplification
- Average 1,000X coverage of Lassa genome
Ancients Origins of Lassa Virus

Andersen et al., Cell (2015)

Kristian Andersen
Scripps
Pardis Sabeti
Harvard/Broad

Ebola virus

Clade IV
Josiah

Clade II
Nig237

Clade III
Nig08

Negative Control

Clinical Sequencing Uncovers Origins and Evolution of Lassa Virus

Highlights
- Lassa virus is a life-threatening pathogen that is endemic in West Africa
- Lassa virus has diverse and ancient origins in Nigeria
- Wild civets from Nigeria and Sierra Leone differ in their transmission efficiency
- The virus evolves within hosts to evade immune-determined selection pressures

Andersen et al., Cell (2015)

Christian Happi
RUN
Metagenomics of Clinical Samples: Identification of Pathogens based on Sequences

Countermeasures
- Vaccines
- Drugs
- Diagnostics

GTACTGACTACGTAGC
GACTGCTGACTGATCG
ATCGATGTATAGCTAC
TAGCTCGCTGCTAGAG
TCGTAGCTAGCTACGTA
TCGATCGTACGTACGTA
CGTAGCATCGATCGTA
CGATGCTAGCTACGAT
GTAGCTAGCTAGCTAC
GTAGCTACGTAGCTAC
ATCGTCGTAGTACGCA
GCCGCCGCTGCCGCCGA
CGATCGCCGCCGATCGG
Febrile samples contain more viral reads

Healthy

- ‘no hit’: 36%
- bacteria: 55%
- viruses: 2%
- eukaryotes: 7%

Febrile

- ‘no hit’: 32%
- bacteria: 48%
- viruses: 8%
- eukaryotes: 12%
<table>
<thead>
<tr>
<th>Virus</th>
<th># of healthy controls</th>
<th># of febrile patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta papilloma virus</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Dengue 3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>GB virus C</td>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>Hepatitis B</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Hepatitis delta virus</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>HIV-1</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Lassa fever virus</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Novel picornavirus</td>
<td>12</td>
<td>54</td>
</tr>
<tr>
<td>Novel rhabdovirus</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>
Discovery of Novel Rhabdoviruses in the Blood of Healthy Individuals from West Africa


© ViralZone 2006
Swiss Institute of Bioinformatics

Tibrovirus
Ephemovirus
Vesicular
Lyssa
Nuclearhabdovirus
Novirhabdovirus
The 2018 clade was more closely related to sequences from other West African countries than to earlier (1946–1991) Nigerian sequences.

- 2018 YFV sequences formed a tightly clustered clade.
Metagenomic analysis reveals Enterovirus B- infant suspected with VHF in Lagos, Nigeria.

- Sample obtained from a child at IKGH, Lagos, Nigeria.
- Metagenomic analysis revealed the presence of Enterovirus B (EV-B).
- Phylogenetic analysis also confirmed CV-B3 as the genome clustered closely with a Coxsackievirus-B3, a serotype of Enterovirus B.
- EV-B is the largest enterovirus species and it causes mild infections but also cause encephalitis, myocarditis, poliomyelitis, acute heart failure and sepsis.
- Neonatal infections with coxsackievirus-B3 are characterized by severe illnesses with myocarditis or meningoencephalitis.
- There have been reported cases of CV-B3 in Nigeria (Faleye et al., 2017) as the virus was found in fecal samples obtained from children <15 years old diagnosed with acute flaccid paralysis.

Mid-point rooted maximum likelihood tree showing relationship between our sequence from this study (coloured blue) and enterovirus reference genomes from the NCBI database.
Next Generation Sequencing for suspected COVID-19 Samples

OVER 70,000 COVID-19 SUSPECTED SAMPLES

TESTING

METAGENOMIC SEQUENCING

Boost testing capacity from 300 to over 5000 per day at ACEGID
The sample of the index case was sent to the ACEGID lab for sequencing on 1st March, 2020.

Using one of the two Illumina MiSeqs in the sequencing platform of ACEGID, we sequenced the sample and obtained a full genome of SARS-CoV-2.

We obtained all HCoV whole genome sequences obtained from human hosts with geographical annotations from GISAID and aligned with the index genome sequence from Nigeria.

The genome clusters with a European clade, consistent with the known travel history of this case (Figure 1).

First SARS-CoV-2 genome sequenced in Nigeria and Africa.

Figure 1: Maximum likelihood tree of SARS-CoV-2 including Nigeria’s index case
Genomics Characterization of the First Rabbit Haemorrhagic Disease Virus in Sub-Saharan Africa

Microbial metagenomic approach uncovers the first rabbit haemorrhagic disease virus genome in Sub-Saharan Africa

EXTENSIVELY DRUG RESISTANT TUBERCULOSIS IN NIGERIA, WEST AFRICA

- Positive Xpert + line probe assay (LPA) drug resistant and multidrug resistant TB cases from clinical samples (n = 2)
- One HIV positive and One HIV negative
- LPA negative to second-line drugs using MTBDRsl
- Samples tested drug resistant and pre-XDR using phenotypic drug susceptibility test (DST)
- Whole genome sequencing using Illumina iSeq100 and bioinformatics analysis confirmed both samples as XDR TB and Beijing lineage of MTBC
- HIV positive + XDR TB patient reported dead whilst HIV negative + XDR TB patient discontinued treatment after adverse side effects to TB medications.

Maximum likelihood phylogeny of 220 MTBC genomes showing drug resistance profiles, geographic regions and lineages. The red arrow shows TB samples from this study.
Metagenomic sequencing characterizes a wide diversity of viruses in field mosquito samples in Nigeria

Emergence and Genomic characterization of first cases of Candida auris in Nigeria and West Africa

- Four different hospitals in Nigeria (2 private ICUs in Lagos, a tertiary hospital in Lagos and another in Ibadan) reported *Candida* bloodstream infections
- WGS done at ACEGID revealed mutations in the ERG11 gene responsible for multidrug resistance in 2 of the isolates
- Phylogenetic analysis also showed that the isolates belong to Clades 1 and 4 of *C. auris*. This suggests at least two independent introduction/emergence into Nigeria
Figure 1A: Proportion of non-LASV assembled genomes of each virus from LASV-positive and LASV-negative samples: Figure 1B

: Frequency of viral genome identification in LASV-positive vs. LASV-negative samples. Asterisks show the significance of a two-sample z-test for difference in proportions. * p < 0.05; **, p < 0.01, *** p < 0.001…
Increased Prevalence of Lassa Fever Virus-Positive Rodents and Diversity of Infected Species Found during Human Lassa Fever Epidemics in Nigeria

Metagenomics and Detection of AMR in Nigeria

- Nasopharyngeal samples from Negative SARS-CoV-2 samples in patients with fever in Ikorodu, Lagos subjected to metagenomic analysis using NextSeq2000 at ACEGID

- More than 56% of the samples had antibiotic resistance markers to Penicillin, Tetracycline, Sulfonamides and Lincosamides

- The samples also show the presence of respiratory viruses such as SARS-CoV-2 and Measles, and also Epstein-Barr virus which primarily spreads through saliva

- Other bacteria pathogens which were present in majority of the samples were Streptococcus pneumoniae and Campylobacter concisus
VGEA: an RNA viral assembly toolkit

Published September 6, 2021  PubMed 34567846

The Bacteria Genome Pipeline (BAGEP): an automated, scalable workflow for bacteria genomes with Snakemake

Published October 27, 2020  PubMed 33194387
Unbiased NGS Matagenomic has significantly advanced our ability to understand the microbial landscape in many part of the world.

Metagenomics has transformed the genomic surveillance and our ability to discover new pathogens.

However, it is vital to develop and implement pan-viral enrichment (removing human/host DNA) approach in order to drastically improve the efficiency of metagenomic sequencing.

Cost reduction is needed in order to sustain surveillance/viral detection and pathogens discovery and clinical applications.

Developing ultrasensitive Bioinformatic pipelines for de novo genomes assembly and taxonomic classification is very crucial.
Acknowledgements

Outstanding Partners
Redeemer's University
Irrua Specialist Teaching Hospital
NCDC
Viral Hemorrhagic Fever Consortium
Broad Institute of Harvard and MIT
Harvard University
Tulane University
UCAD, Senegal
KGH, Sierra Leone
NPHI, Liberia
Zalgen
University of Cambridge
FETHA
FMC-Owo

Generous Funders

outbreaks genomic response team
More Information

- www.run.edu.ng
- www.acegid.org

Or e-mail
Info.acegid@run.edu.ng