Public Health Alliance for Genomic Epidemiology

PHA4GE Newsletter

August 2021

Editorial

As we continue as a global collective to navigate the COVID-19 pandemic, it is certainly a privilege for the PHA4GE consortium to share some practical implementation projects that have been initiated since our last communication. Our concerted effort to implement data standards in public health laboratories was accelerated by funding from the Bill and Melinda Gates Foundation to support ten research teams in Africa and Southeast Asia. These teams partner with public health institutes to assess and implement data standards and reproducible bioinformatics pipelines to achieve a timely public health response for the control of disease outbreaks. In this newsletter we introduce each of these 10 teams.

The Ethics and Data Sharing working has gained momentum. Among the projects that are highlighted by Anja Bedeker is a meta-data standard for ethical research data. We look forward to wider participation as we grapple with the very real challenge of sharing pathogen data in the context of the COVID-19 pandemic.

Jamie Southgate reports on the Bioinformatics pipelines working group's recent publication. The group released a guidance document that defines the major challenges faced in public health to integrate technologies in response to SARS-COV-2 analysis and to highlight various open-source resources available in the public health space.

In this newsletter we get to know Dr Josefina Campos, Director of "National Center of Genomics and Bioinformatics

-ANLIS 'Dr Carlos G Malbrán", which forms an integral part of the National Administration of laboratories and for Health Institutes in Argentina. We get insight into the genomics surveillance work in Argentina and how it interfaces with the broader Latin American bioinformatics public health initiatives.

Rangarirai Matima discusses organisational resilience and its associated enablers in the context of healthcare systems that face public health threats.

Alan Christoffels

PHA4GE sub-awards Announcement

Concerted effort among different collaborators is required to generate a public health response for the control of disease outbreaks. Among the collaborators, public health laboratories play a major role as they answer the "What is causing this disease?" question. This is through sequencing samples to detect pathogens of concern. However, sequencing laboratories use different bioinformatics platforms to get this data. This creates a common data sharing challenge – the lack of interoperability of datasets and databases.

To minimise such a challenge, our consortium's Data Structure Working Group developed two standardised open-source tools: "***hAMRonization***" a gene detection output for anti-microbial resistance [**(AMR) - quick facts**] and a contextual data specification for severe acute respiratory syndrome coronavirus 2 [**(SARS-CoV-2) -quick facts**]. With support from other Working Groups in the consortium, there is uptake and utilisation of these specifications across

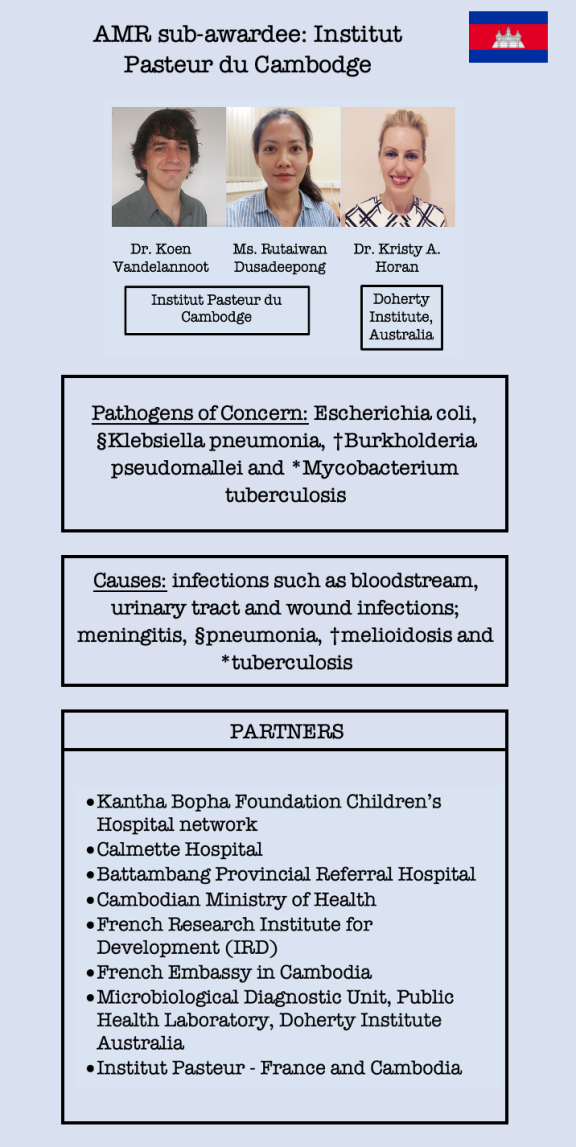
different public health laboratories globally.

Notably, in select low to middle income countries (LMICs), funds from the Bill and Melinda Gates Foundation have become instrumental to further utilisation of this specification and other advancements in bioinformatics for public health. Ten research teams will be partnering with national public health laboratories to implement standardized bioinformatics practices, pipelines, and data structures in either AMR or SARS-CoV-2 sequencing.

In this edition, we share the profiles of the recipients of these sub-awards.

AMR theme recipients are from **Cambodia**, **Malaysia** and **Nigeria**. Teams will build genomic AMR surveillance capacity in their countries and regionally through: 1) developing data structures for AMR pathogens relevant to their region 2) implementing bioinformatic pipelines to analyse genomic AMR data, where required 3) AMR data exchange and comparison across partner laboratories and beyond 4) conducting training sessions coupled with support visits to each sentinel site on next-generation sequencing (NGS)-(applicable to Nigerian context only).

AMR theme sub-awardees:



AMR sub-awardee: Universiti Kebangsaan

Assoc. Prof. Neoh Hui Min | Prof. Sheila Nathan | Dr. Lam Su Datt | Dr. Tengku Zetty Maztura Binti Tengku Jamaluddin

National University of Malaysia (Universiti Kebangsaan Malaysia)

Mr. Mia Yang Ang

University of Tokyo, Japan

Dr. Sabrina Di Gregorio

Universidad de Buenos Aires, Argentina

Pathogens of Concern: Methicillin-resistant Staphylococcus aureus (MRSA)

Causes: skin infections that may progress to the bloodstream, bones and joints

PARTNERS

- National University of Malaysia (UKM) [Medical Molecular Biology Institute, Faculty of Medicine and Canselor Tuanku Muhriz (HCTM) teaching Hospital]
- Universiti Putra Malaysia (UPM)
- Institute for Medical Research (IMR)
- Ministry of Health, Malaysia
- Malaysia Laboratory of Research on Bacterial Resistance to Antibiotics
- Division of Microbiology, University of Buenos Aires, Argentina

AMR sub-awardee: Global Health Research Unit for Genomic Surveillance of Antimicrobial Resistance

Prof. Iruka N. Okeke | Dr. Ayorinde O. Afolayan | Ms. Ifeoluwa Akintayo | Mr. Erikson. E. Odih

Ms. Faith Oni | Mr. Rotimi Dada | Dr. Odion Ikhimiukor

Global Health Research Unit for Genomic Surveillance of Antimicrobial Resistance, University of Ibadan, Nigeria

Pathogens of Concern: World Health Organization priority pathogens including (but not limited to), Escherichia coli, Acinetobacter baumannii, §Klebsiella pneumoniae and ¤Salmonella

Causes: infections such as bloodstream, urinary tract and wound infections; §pneumonia, and ¤intestinal tract infection

PARTNERS

- Nigeria Centre for Disease Control
- Ten public health laboratories

*a mix of sentinel surveillance hospital laboratories in South-west Nigeria and private laboratories in Lagos

For the SARS-CoV-2 theme, the recipients are from **Fiji, Kenya**, **Malawi**, **Nigeria**, **Pakistan**, **The Gambia** and **Zambia**. Teams will focus on strengthening SARS-CoV-2 disease surveillance and where possible, identification of variants of concern in their regions and globally through: 1) establishing sequencing protocols that can generate high-quality SARS-CoV-2 sequence data 2) training and capacity development in bioinformatics -protocol and automated bioinformatics pipeline development, sequencing and data analysis of bioinformatics workflows 3) presenting analysed data among partners in a form suitable for public health utility and inform public health policy 4) uploading SARS-CoV-2 sequence data and relevant metadata onto international genomic databases 5) growing regional scientific capacity to tackle the ongoing and future pandemics by establishing new and strengthen existing partnerships.

SARS-CoV-2 sub-awardee: Fiji Centre for Disease Control


Dr. Aalisha Sahukhan | Ms. Shalini P. Singh

Fiji Centre for Disease Control



Prof. Benjamin Howden | Dr. Kristy A. Horan | Ms. Tuyet Hoang

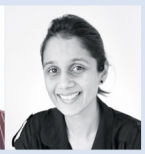
Doherty Institute, Australia

PARTNERS

- Fiji Ministry of Health and Medical Services (MoHMS)
- Fiji Centre for Disease Control (CDC) [Molecular Laboratory]
- Microbiological Diagnostic Unit, Public Health Laboratory, Doherty Institute, Australia

SARS-CoV-2 sub-awardee: International Livestock Research Institute





Dr. Jean-Baka Domelevo Entfellner | Dr. Samuel Oyola | Dr. Sonal Henson

International Livestock Research Institute, Kenya

PARTNER

Ministry of Health Kenya

SARS-CoV-2 sub-awardee: College of Medicine

Ass. Prof. Arox Kamng'ona | Dr. Khuzwayo Jere | Dr. Benjamin Kumwenda | Dr. Philip Ashton

College of Medicine, University of Malawi

PARTNERS

- Malawi Liverpool Wellcome Trust
- Ministry of Health Malawi
- Queen Elizabeth Central Hospital
- Public Health Institute of Malawi

SARS-CoV-2 sub-awardee: Aga Khan University

Ass. Prof. Waqasuddin Khan | Ass. Prof. Muhammad Imran Nisar | Ass. Prof. Fyezah Jehan

Ms. Samiah Kanwar

Mr. Furqan Kabir

Aga Khan University, Pakistan

PARTNERS

- Chan Zuckerberg Initiative
- Aga Khan University – Department of Pediatrics and Child Health
- National Institute of Health
- Ministry of Health, Pakistan

SARS-CoV-2 sub-awardee: Africa Centre of Excellence for Genomics of Infectious Diseases

Prof. Christian T. Happi | Dr. Idowu B. Olawoye | Dr. Paul E. Oluniyi

Africa Centre of Excellence for Genomics of Infectious Diseases (ACEGID), Redeemer's University, Nigeria

PARTNERS

- Nigeria Centre for Disease Control
- Africa Centre for Disease Control,
- African Pathogen Genomics Initiative
- Nigerian Ministry of Health

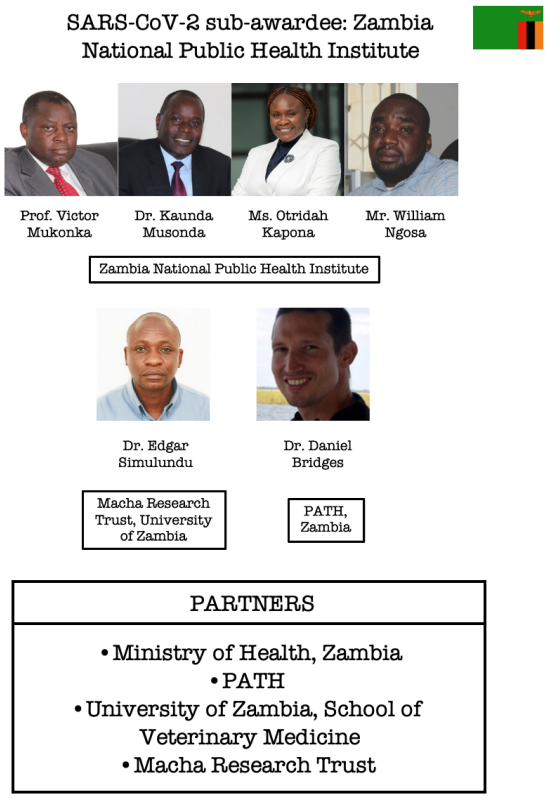
SARS-CoV-2 sub-awardee: Medical Research Unit, The Gambia

Dr. Abdul Karim Sesay | Mr. Abdoulie Kanteh

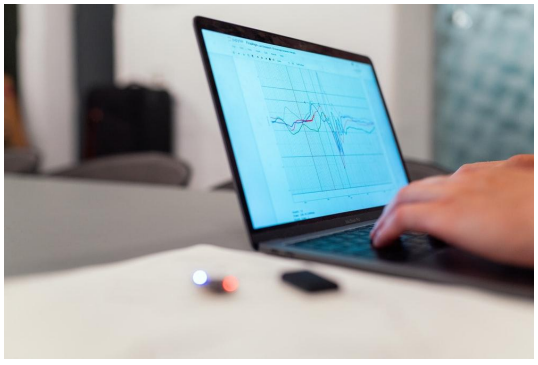
Medical Research Unit, The Gambia (London School of Hygiene and Tropical Medicine)

PARTNERS

- National Public Health Laboratory, The Gambia
- University of Cheikh Anta Diop, Senegal
- Nanaro Clinical Research Site, Burkina Faso
- Universidade Jean Plaget, Guinea-Bissau
- National Institute of Medical Research, Nigeria



Bioinformatics Solutions for SARS-CoV-2 Genomic Analysis



Genomic analysis of SARS-CoV-2 has made a profound impact in support of the global response to the pandemic. The ability to carry out analysis on samples is increasingly being considered as a crucial functionality of public health labs. However, many labs are faced with the daunting task of starting an entire bioinformatics program and integrating the appropriate tools and software from the ground up. The benefits these technologies bring to public health are invaluable.

In an effort to bridge the gap in integrating these technologies for SARS-CoV-2 (SC2) analysis, PHA4GE Pipelines and Visualization Working Group members collaborated on a **guidance document** to define the major challenges and to highlight various open source resources that have emerged from the public health community. Conceptualizing the challenges faced aided in identifying and accentuating the major open access

and open source resources available in the public health space.

One of the major challenges identified by the Working Group is the generation of census genome assemblies from PCR tiling NGS Data. Tiled amplification sequencing through the **ARTIC protocols** for example is one of the most commonly adopted approaches for generating SC2 sequencing data. These sequencing experiments generate thousands of amplicon reads that represent fragments of the original SC2 genome present in the sample. The raw data that these labs generate are often amplicon reads. As a result, one of the initial bioinformatics demands that laboratories face is the assembly of PCR tiling Next-generation sequencing (NGS) data into a contiguous SC2 genome. The generation of consensus genomes is incredibly powerful as it allows for further downstream analysis such as lineage typing and genomic epidemiology studies that help to inform public health decision making.

Working Group members noted across multiple public health laboratories, was the sharing of SARS-CoV-2 sequence data. The sharing of sample read and assembly data through internationally accessible databases allows insights to be drawn about how the virus is spreading and mutating across the globe. In making the data available to international researchers, strong minded public health decisions can be made; nonetheless preparing and submitting data to these repositories can be a challenge in itself.

The screening for Variants of Concern (VOC) or in essence making lineage and clade assignments for sequenced SC2 samples bears crucial influence on the decisions made by public health officials. Thus the invitation to accurately and reliably screen for VOC's such as B.1.617.2 (Delta) is a critical component to the Bioinformatics analysis of SC2 genomes carried out by public health laboratories.

Genetic relatedness as inferred through performing phylogenetic analysis on collections of SC2 samples can be a powerful proxy for epidemiological associations that can help resolve transmission networks, enable real-time surveillance, provide insights on genetic variance over time and support local outbreak interventions. Gaining access to the Bioinformatics Solutions for SARS-CoV-2 phylogenetic analysis can greatly benefit public health efforts.

Particular attention was made to demonstrate open access and open source solutions to these tasks to reduce the barrier to access and can be assessed thoroughly by the greater community of public health bioinformatics practitioners. We knew that as soon as we made this resource document available, that necessary modifications were inevitable to account for new software, resources and other approaches released by the public health bioinformatics community. With this in mind, the document is hosted on a public **GitHub page** to enable timely updates from the working group but also continuous community contributions. A special thank you to all Working Group members and external persons who participated in the creation of the document. It is worth noting that the document is a reflection of the opinions of our Working Group and to enhance its value we gladly welcome and encourage external collaboration. Additions to the document can be made via raised issues/pull requests or emailing us at the Working Group (**pipelines-visualisation@pha4ge.org**)

On the 24th of June, the Working Group hosted Dawn Roellig and Jillann Hagey from the US CDC's Technical Outreach and Assistance for States Team (TOAST) who presented their Menu to assist labs starting out with SC2 sequencing. The Menus cover wet-lab (library preparation) all the way to submitting data to public repositories whilst highlighting the various workflow

options available for Illumina and Oxford Nanopore.

Forthcoming, the Working Group is making great progress on a Quality Control (QC) Solutions for SARS-CoV-2 genomic analysis collaborative document which will be released in the coming weeks. NGS has expanded the approach of genomic analysis for pathogen surveillance systems, While the demand for NGS continues to grow, the quality of NGS sequencing data can be affected by library preparation and sequencing processes, systematic variation in quality scores across sequence reads, biases in sequencing due to base composition, and less-than optimal library fragment sizes and indexes. The collaborative document aims to assist in defining the QC challenges for SC2 genomic analysis and suggest QC systems solutions to address them. Keep a look out on the PHA4GE social media pages and website for further announcements.

In the upcoming quarter, the Working Group has set its sights on discussing validation sets solutions for SC2 genomic analysis; Informing public health action through genomic visualization and dashboarding; and supporting the research software developer arena in assisting in collaboration, containerization and showcasing best practices.

In conclusion, if any of the topics mentioned in this update may be of particular interest, please feel free to join the Working Group to participate in discussions at our meetings or on the PHA4GE Slack Channel. New contributions are immensely welcomed.

~ by Jamie Southgate, on behalf of the Bioinformatics Pipelines and Data Visualization Working Group

Ethics and Data Sharing Working Group update

The **PHA4GE Ethics and Data Sharing Platform** continues to grow as more people are joining and participating in the community.

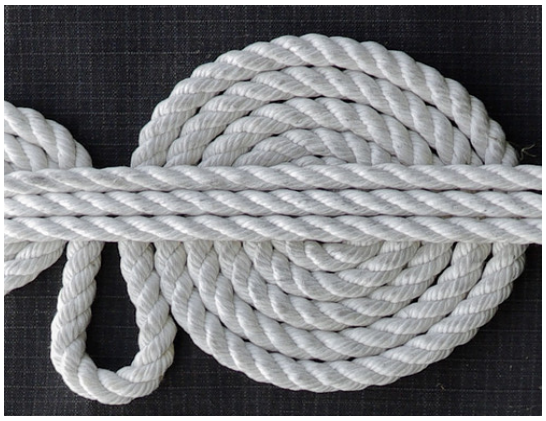
Several projects have kicked off from the platform (e.g. building frameworks and undertaking a systematic review on risks and benefits in health genomics studies in Africa, and a project to define a meta-data standard for ethical research data), which has led to these projects including collaborators from different countries and continents and different professional, educational and disciplinary backgrounds. This has made the sharing of ideas, expertise and the content of the projects richer, more diverse and more inclusive.

Community members are also reaching out asking for advice and sourcing the opinions and expertise of other community members on different topics, for example ethical issues that should be considered when sharing pathogen sequence data.

We invite anyone interested to join us on this interactive platform, there is a short process to create your user account and then you can join our conversations and initiate or participate in projects about ethics and data-sharing.

~by Anja Bedeker, on behalf of the PHA4GE Ethics and Data Sharing Working Group

OPINION: Organisational resilience in the context of public health threats



Organisational resilience is championed across various sectors such as urban planning, ecology, extractive industries, and healthcare for a **system's ability to continue to meet its objectives in the face of challenges.**

With regards to resilient cities, cities embody productivity and are seen **as functional elements of the global political economy.** Various ideas and recommendations are highlighted as **criteria to resilient urban planning** and directions are given for future research in **sub-Saharan Africa** and resilient urban-development in **Asia**, **America**, **Central Europe**, **Western Europe** and **Oceania**.

Climate change adaptation and resilience is vaunted for coastal **regions, dryland systems** and **semiarid regions, deltas, glacier- and snowpack-dependent river basins** across Africa and Asia. This is particularly true to maintain balance in ecosystems and especially for an important element for human and animal survival – food production. For instance, to enhance resilient African agricultural systems, **climate services ought to be incorporated into national policies to avert the threats of climate change** and there must be

an increased **support for production of African indigenous vegetables**.

The extractive industry also highlights how stakeholder engagement and buy-in from companies and the surrounding communities are key in developing and implementing sustainable plans in **South American** and **African** contexts.

But what is organisational resilience in healthcare? In the context of emerging and known public health threats, how is resilience fostered?

There is no universal definition of resilience in empirical health research. There are various connotations due to the varied meanings that are ascribed to the term. This may be attributed to the fact that the term "resilience engineering" was borrowed from safety science and translated to healthcare. **Resilience engineering** is concerned with organisational processes that enable a team or unit to adapt successfully. Three premises underpinning resilience engineering include: *'resilience at organizational level'*; *'variability in the environment creates the need for adjustment'*; and '*the need for adjustment and the ubiquity of adaptation means that there is a difference between work as imagined (WAI) in protocols, procedures and targets and work as done (WAD) in practice'*.

Some review papers in healthcare have illustrated the evolution of the term resilience engineering by using terms such as "resilient healthcare", "planned, adaptive and everyday resilience". **Resilient healthcare** is the application of the concepts and methods of resilience engineering in the healthcare field, specifically regarding patient safety. **Planned resilience** is a function of planning for and preparing for future crises; **adaptive resilience** speaks to adapting to chronic stresses and acute shocks and **everyday resilience** is the

resilience of health systems to routine and chronic stress.

Public health threats, that fall in different classes due to different characteristics and attributions, become ideal to assess resilience in healthcare. Looking at public health emergencies of international concern (PHEIC), **a PHEIC is a formal declaration by the World Health Organization** of '*an extraordinary event which is determined to constitute a public health risk to other States through the international spread of disease and to potentially require a coordinated international response'*. This includes events such as H1N1 influenza, Ebola, poliomyelitis, Zika and COVID-19. Antimicrobial resistance (AMR) is another classification of a public health threat, which **is one of the top 10 global threats facing humanity** *'as the misuse and overuse of antimicrobials are the main drivers in the development of drug-resistant pathogens' and 'without effective antimicrobials, the success of modern medicine in treating infections, including during major surgery and cancer chemotherapy, would be at increased risk.'*

Zooming into Ebola, COVID-19 and AMR, challenges and opportunities in fostering organisational resilience are reported. **Prior to the 2014–2015 Ebola outbreak, infection prevention and control (IPC) activities in Liberian healthcare facilities were basic** with *'neither national IPC guidance nor dedicated staff at government and health facility levels to ensure implementation of best practices'.* **A ripple effect of unintended consequences also spread across other sectors** of economic significance such as agriculture and regional trade that resulted in food shortages and a rise in unemployment, respectively. ***'Substantial external financing' was seen as a need 'to build stronger national and subnational health systems'*** and **immense financial support came through different channels** such as the International Monetary Fund, World Bank, European

Union and African Development Bank. The **deployment of international medical experts** to support local healthcare facilities, **resilient leadership through the influence of political and cultural leaders** over the public to adapt and cooperate to national and international guidance in controlling the epidemic and **further training of healthcare workers in IPC and relaying of standardized messaging on Ebola to the public** were contributory factors to the success of the concerted national and global efforts in containing the epidemic.

Resilient leadership in the context of COVID-19 has been described as "paradox work" as leaders face *'contradictory tensions'* and *'apparent contradictions'* that push them to be *'reactive and adaptive'* in ever changing circumstances. In the early months of the COVID-19 pandemic, **health systems in Asia** that have *'appropriate containment measures and governance structures; took steps to support health-care delivery and financing; and developed and implemented plans and management structures'* but were uncertain if they could continue to function properly with the health emergency. **A downward spiral of sub-Saharan economies post COVID-19 is anticipated** if '*urgent and timely domestic and international interventions'* do not exist.

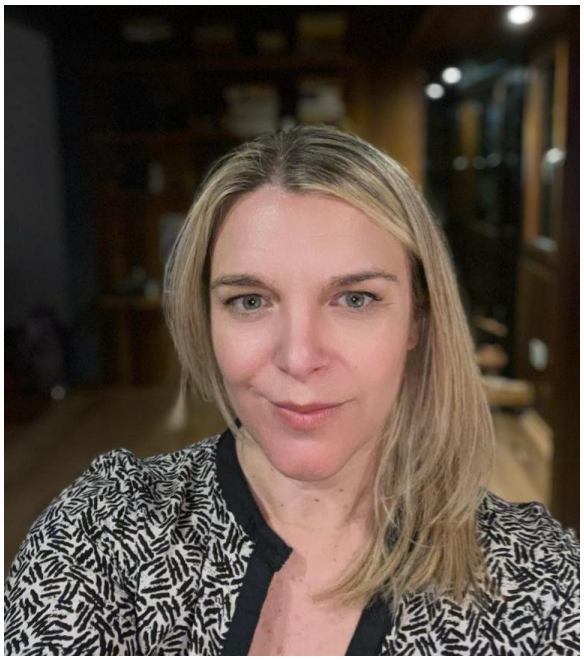
Exercise of corporate social responsibility for COVID-19 is one mechanism to alleviate the impact of COVID-19. For example, some private companies in **Spain** donated *'health, food, and research funds, and temporarily* transferred *the necessary infrastructure, among other action'.* In the **Philippines**, such efforts have been proposed where private companies could purchase COVID-19 vaccines for their employees and *'the government can remove taxes on these companies for this transaction only'* as a way of strengthening the national vaccination program.

Fears of unintentional exacerbating AMR are also highlighted in managing COVID-19, as treating physicians may prescribe antibiotics more often to control suspected underlying bacterial infections in patients with COVID-19 infection. Sustainable and resilient antimicrobial stewardship (AMS) programmes may be built through promotion of capacity building in AMR, reciprocal and long-term partnerships between high-income and low- and middle-income countries and adequate funding.

With the cases described above, the availability of resources such as training and financial resources, coupled with multi-sectoral collaboration and resilient leadership - become major enablers of organisational resilience especially when healthcare systems face public health threats. These strengths may be nurtured and reinforced especially across **African health systems that are historically labelled as 'fragile'** and beyond.

~by Rangarirai Matima

Community Profile - Dr. Josefina Campos



In this quarter's edition of the PHA4GE newsletter, we shine a spotlight on Dr. Josefina Campos, the Director of the National Center of Genomics and Bioinformatics -ANLIS 'Dr Carlos G Malbrán', which forms an integral part of the National Administration of

laboratories and for Health Institutes in Argentina. Dr. Campos is also a member of the PHA4GE Data Structures Working Group. The discussion was captured during an interview between Dr. Emma Griffiths (Chair, Data Structures WG) and Dr. Campos in late July.

Dr. Campos, what is your current affiliation and position?

I am the Director of the Unidad Operativa Centro Nacional de Genómica y Bioinformática at ANLIS "Dr Carlos G. Malbrán" in Buenos Aires, Argentina. It is a new, dedicated Genomics and Bioinformatics center. The lab used to be part of another Institute, but due to the demonstrated importance of genomics and bioinformatics during the COVID-19 pandemic, the lab has recently expanded and is now its own institute.

Wow, that's fantastic! Are you only sequencing SARS-CoV-2 in the lab?

We are definitely sequencing SARS-CoV-2 genomes from across the country, but our scope is quite broad. We also sequence foodborne pathogens, and also incorporate environmental genomics because linking environmental, animal and human sources of pathogens using a One Health approach is incredibly powerful for investigating and responding to infectious disease. We also have an interest in antimicrobial resistance, human genomics, TB, Human Papillomavirus, Neisseria meningitidis and Neisseria gonorrhoeae, and other areas which we work hand by hand with the National Reference Laboratories that do the national surveillance.
But we don't just do whole genome sequencing, we use a range of genetics and bioinformatics techniques.

How big is the lab?

Before COVID, we had 4 people in the lab - myself and three staff members. Currently we have seven, and by the end of the year we aim to keep incorporating members. We anticipate that the lab will continue to grow as well next year. We are in a dedicated BioSafety level 2 lab for our genomics work. We have a NovaSeq, 2 MisSeqs, and 4 robots to help with sample prep and automation.

Have genomics and bioinformatics been a public health priority for a long time in Argentina? And how is Argentina networked to other countries in Latin America?

I would say our journey first began when we got involved with PulseNet Latin America and the Caribbean (PulseNet LAC). PulseNet were the pioneers in building a foodborne pathogen molecular surveillance network. I attended a course on genomics surveillance as an instructor for pulse-field gel electrophoresis (PFGE) at the Wellcome Trust Advanced courses in 2013 and knew nothing about it, and had to read as much as I could before the course. Before that, I was put in charge of training labs to perform PFGE all across Latin America. I've visited most countries through that training program. Then things started to move fast when PulseNet shifted to genomic surveillance. In 2016, we built the sequencing lab with basically myself, one tech and a MiSeq. I worked with Ruth Timme (US FDA) to learn sequencing and analysis techniques so we could participate in **GenomeTrakr**.

We work in collaboration of a number of networks lead by PAHO now across Latin America which include **PulseNet**, as well as La Red Latinoamericana de Vigilancia de la Resistencia a los Antimicrobianos (**ReLAVRA**, or The Latin American and Caribbean Network for Antimicrobial Resistance Surveillance in English). Currently, the network is one of the oldest, and largest regional AMR surveillance networks in the world. This network was started by the WHO/PAHO regional office and partnering member states in 1996 and now includes 19 national reference labs, in 20-member states, that regularly report antimicrobial susceptibility information from sentinel labs covering a wide range of foodborne, nosocomial and community-acquired pathogens e.g. Salmonella, Shigella, V. cholerae, E. coli, Staphylococcus aureus, Acinetobacter baumannii, Klebsiella spp. Pseudomonas aeruginosa, Enterobacteriaceae, Enterococci spp., S. pneumoniae, N. meningitidis, and

Haemophilus influenzae. We are now incorporating genomic surveillance into the region and have joint meetings between ReLAVRA and PulseNet to share protocols and solve common issues. However, outside of PulseNet, ReLAVRA also focuses on identifying AMR genes and mutations in pathogens, as well as which databases are the most appropriate for our use cases. We have also participated in the recently formed **Regional Genomic Surveillance Network** that sequences and tracks COVID-19 virus variants throughout Latin America and the Caribbean, which arose from the existing influenza network.

What have been some of your greatest successes doing this work?

Definitely capacity building - using different strategies to bring sequencing to different Latin American countries. You know, not very long ago - maybe 3-4 years ago - no one knew what genomics was in the training courses. Now we are targeting **how** to implement genomics (**not if** we should implement it), and we are doing it at a level on par with "high income" countries. Previously we were always trained by others, now we are at the forefront which I think puts us in a stronger, more equitable position in global conversations. Also, we are very proud of our collaborations - the consensus building we have been able to achieve (both within Argentina where our sequencing lab works hand-in-hand with the national reference lab, as well as across Latin America) in a very short time. The Networks have been a very powerful tool for this.

What are some of the challenges you have faced in setting up genomics and bioinformatics labs and surveillance networks?

Implementing genomics and bioinformatics is not as simple as giving someone a tech and a sequencer, you have to think through the process from end-to-end. This was really brought home to me when I used to travel to different labs to teach PFGE training. Sometimes a lab might not have something simple like a shaker/water bath, which would mean we would have

to troubleshoot. Now, when we are working on implementing genomics, sometimes labs don't have enough cold storage for samples, or have internet issues or unreliable electricity supplies, so that data transfers get interrupted. There are so many different contexts, so many different realities for people that it is hard to think globally - to think outside your own experience. That's why it is so important to listen and to hear different voices when you are creating networks and implementing new technology in public health. So when we are seeking funding, it's not just about paying for the staff or the sequencer, you have to think about procuring reagents (the cost and whether the supply is reliable), the freezers and storage containment, the Qubit, the information management system, the compute infrastructure, and so on.

Because of COVID, there has been a lot of change in the last year and a half. We have managed to build a lot of capacity in the region, but now we must consider how things will change in the future - how genomics can be used beyond the pandemic. We need to prioritize which pathogens to target with genomic surveillance. We need to identify new funds. As many of these initiatives began as research projects, we need to think about when it's useful to use genomics because it's very difficult to do it for everything. We are asking questions about how to evolve the culture of genomics to make it work for public health. Genomics is not just about data generation, analysis and interpretation, it also has consequences for policy makers. Every research publication we achieve helps with public policy changes. For example, our genomics work with cholera with the Sanger Institute demonstrated that we could distinguish between endemic and pandemic lineages, which had direct intervention implications i.e. endemic infections could be dealt with with regular procedures but pandemic lineages needed stronger interventions.

Pursuing funding can be a challenge for researchers in public health in LMICs. Often, if we utilize funding pathways used by HICs, we are co-PIs in

international grants instead of PIs. That can be because of issues of trust regarding the disbursement of funds from international funders. So we need to figure out what are the best opportunities, as we need to collaborate with the right people. Also, it can be difficult to promote your work because we can't afford to go to high profile conferences, we are not invited to expert committees, if English is not our mother tongue it can be more difficult to publish in different journals. This all contributes to making us less visible.

How does your career work relate to the work you do in PHA4GE?

The PHA4GE work intersects with everything. I'm a PI on a new grant from the Wellcome Trust for strengthening COVID genomic surveillance in Latin America. It's a one year grant, but it positions us to expand from SARS-CoV-2 to other pathogens. We will be able to start Nanopore sequencing in regions that lack capacity in order to increase their representation, and build up the network among the centres that are already generating data. Our goal is generating consensus among Latin American countries in reporting, interpreting results, and how we share/upload to public repositories. We will work together with PAHO as the link for this work, but we will be implementing the **PHA4GE SARS-CoV-2 contextual data specification** among the network to better enable harmonization across labs. We have also been piloting **PHA4GE's hAMRonization tool** together with ReLAVRA and PulseNet labs using genomics surveillance for AMR, which helps to harmonize the results of different AMR gene detection tools and provides interactive and tabular reports. In that pilot, we translated the PHA4GE information package for policy makers and installation instructions for bioinformaticians into Spanish. I think we will be including a story about that work on the PHA4GE website shortly. We will try to work on translation of other materials with PHA4GE in the future as well.

What is next for the Data Structures WG?

The Data Structures WG has plans for a lot of exciting new projects in the near future. To complement the AMR gene detection harmonization tool, we will be expanding its capabilities to harmonizing information about point

mutations, which will include expanding the data specification as well as parsers and the harmonized reporting. We are also working on potential solutions for harmonizing different AMR databases. I'm very excited about a brand new project we have been discussing about using the Data Use Ontology to add attributes to metadata shared with public repositories to create a 2-way street of communication between data providers and data consumers. The DUO tags can enable data providers to include information in their submissions about data attribution, collaborations, data use, and more.

Photos courtesy of the ANLIS"Dr Carlos G. Malbrán".

For more information about the activities of the Data Structures Working Group, **click this link**.

Interested in being featured in our newsletter?

Send an email to communications@pha4ge.org

Events

We'd like to share some of the other upcoming events in our community.

Global Alliance for Genomics & Health Collaborate. Innovate. Accelerate.	**GA4GH 9th Plenary Meeting** 28 - 29 September 2021 **Learn More**
ASTMH #Courage #Compassion #Culture 2021 Annual Meeting November 17–21	**ASTMH 2021 Annual Meeting** 17 - 21 November 2021 **Learn More**

Web edition: bit.ly/3bbnxZm

Website: pha4ge.org

Twitter: @pha4ge

Facebook: facebook.com/pha4ge