

Eri Togami et al.

The Global Virome Project

This online first version has been peer-reviewed, accepted and edited,  
but not formatted and finalized with corrections from authors and proofreaders.

## The Global Virome Project

Dennis Carroll,<sup>a</sup> Brooke Watson,<sup>b</sup> Eri Togami,<sup>c</sup> Peter Daszak,<sup>b</sup> Jonna AK Mazet,<sup>c</sup> Cara Chrisman,<sup>a</sup> Eddy Rubin,<sup>d</sup> Nathan Wolfe,<sup>d</sup> Carlos Morel,<sup>e</sup> George Gao,<sup>f</sup> Gian Luca Burci,<sup>g</sup> Keiji Fukuda,<sup>h</sup> Prasert Auewarakul<sup>j</sup> & Oyewale Tomori<sup>k</sup>

<sup>a</sup> Pandemic Influenza and other Emerging Threats Unit, Bureau for Global Health, United States Agency for International Development, Washington DC, United States of America (USA).

<sup>b</sup> EcoHealth Alliance, New York, USA.

<sup>c</sup> One Health Institute, School of Veterinary Medicine, University of California, Davis, California, USA.

<sup>d</sup> Metabiota, San Francisco, California, USA.

<sup>e</sup> Center for Technological Development in Health, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil.

<sup>f</sup> Institute of Microbiology of Chinese Academy of Sciences, Beijing, China.

<sup>g</sup> International Law Department, Graduate Institute of International and Development Studies, Geneva, Switzerland.

<sup>h</sup> School of Public Health, Hong Kong University, Hong Kong, China.

<sup>j</sup> Department of Microbiology, Faculty of Medicine, Siriraj Hospital, Mahidol University, Bangkok, Thailand.

<sup>k</sup> Nigerian Academy of Science, Lagos State, Nigeria.

Correspondence to Eri Togami (email: etogami@ucdavis.edu).

*(Submitted: 20 October 2017 – Revised version received: 4 December 2017 – Accepted: 7 February 2018 – Published online: 5 March 2018)*

At the Prince Mahidol Awards Conference on 30 January 2018 in Bangkok, Thailand, policy- and decision-makers, experts, researchers, donors and private sector representatives from across the globe came together to introduce and explore the dynamics of the Global Virome Project.

The project is an innovative 10-year proposed partnership to develop a global atlas of most of the planet's naturally occurring potentially zoonotic viruses. The project aims to transform the study of emerging diseases by building an unprecedented global atlas of viruses in their ecological contexts. This foundation of information on viral sequences, geographic ranges and host distributions can drive the development of prevention efforts against future threats. This international alliance will connect the next generation of scientists, build capacity at hotspots of viral emergence and promote equitable access to data and strategies to prevent epidemics.

Despite the human and economic impact of viral epidemics, the world is not well enough prepared for the next emerging viral epidemic. Global trends indicate that new microbial threats will continue to emerge at an accelerating rate, driven by our growing population, expanded travel and trade networks, and human encroachment into wildlife habitat.<sup>1</sup> Most emerging viruses are zoonotic, that is, transferred between vertebrates and humans.<sup>2</sup> Nearly all zoonoses originate in mammalian or avian hosts;<sup>3-5</sup> for example, the type 1 human immunodeficiency virus (HIV-1) originated from chimpanzees and Ebola virus from bats.<sup>6</sup> Estimations show that there are more than 1.5 million mammalian and waterfowl viruses, spanning across 25 viral families.<sup>4</sup> Compared to the more than 260 viruses known in humans,<sup>7</sup> the unknown viruses represent 99.9% of potential zoonoses. These viruses usually remain undetected until they cause disease in humans.

Discovering and characterizing viruses in wildlife reservoirs is economically and technologically challenging. However, recent initiatives, including the PREDICT project of the United States Agency for International Development's Emerging Pandemic Threats programme, have shown that systematic viral discovery, even in countries with limited laboratory infrastructure, is feasible.<sup>8,9</sup> Previous studies have identified mammalian species,<sup>10</sup> geographic regions<sup>11</sup> and zoonotic viral transmission pathways<sup>12,13</sup> with increased risk of zoonotic disease emergence. These data enable targeting viral discovery in wildlife to expand our knowledge of likely zoonotic agents with a high potential for spillover to people. The PREDICT project has already discovered over 1000 viruses, including novel Severe Acute Respiratory Syndrome (SARS)-like coronaviruses that can infect human cells.<sup>14</sup>

The Global Virome Project seeks to significantly expand the scale of targeted viral discovery over a decade-long sampling and laboratory testing period.<sup>4</sup> An international consortium of leading disease ecologists, public health practitioners, veterinarians, epidemiologists, biologists and laboratory scientists designed the project.<sup>4</sup> The project's working groups of ecologists, epidemiological modellers and field biologists will select sampling sites and species that harbour the greatest number of unknown zoonoses,<sup>7</sup> and researchers will systematically collect and characterize viruses and their associated metadata in these areas. Protocols for the project's implementation, including training, sampling, specimen handling, laboratory testing, reporting and data management are being developed by the PREDICT project, which has been operating for eight years in over 35 countries.<sup>8,9</sup> The Global Virome Project will

operate as a federation of national and regional projects led by in-country researchers, who are in turn connected to a global hub that provides standardized protocols and monitors progress.

The Global Virome Project seeks to identify the majority of unknown viral diversity, catalogue the ecological conditions at sampling sites, and collect metadata that can be used to analyse the risk of viral spillover into humans. These data will be housed in an open-access database available to the global health community. Such a data set can be used to develop, train and test machine learning algorithms to identify patterns among viruses, classify traits that are more common in zoonoses than in commensal viruses, and predict which viruses have an increased risk emerging in humans.

To ensure equitable sharing of benefits from this project, an ethical, legal and social implications working group has been an integral part of the Global Virome Project since its inception. All research conducted as a part of the Global Virome Project will hold to ethical standards that ensure sharing, including authorship and intellectual property. Central to the ethos of the Global Virome Project is the commitment to building scientific and response capacity in the areas that need it most.

## Building capacity

Novel viruses usually emerge in regions where dense human populations and biodiversity intersect.<sup>11</sup> However, these areas often have limited laboratory, surveillance and health system infrastructure, which delays detecting emerging pathogens and preventing their potential to become pandemic.<sup>15</sup> Thus, to be effective, pandemic prevention should take place at the source of viral spillover events – before they spread regionally.<sup>16</sup> The Global Virome Project seeks to make data on novel potential zoonotic viruses available to public health agencies that face undiagnosed illness in humans and animals. The project has the potential to benefit human and animal health through broad measures, such as: (i) using polymerase chain reaction assays for an expanded diversity of potentially zoonotic viral families to shorten the time between outbreak detection and pathogen identification; (ii) strengthening global epidemic preparedness through investment and training in epidemiological surveillance, field biology, laboratory techniques and biosafety; (iii) identifying high-risk pathogens in wildlife populations that have high contact with people (for example hunted species and peridomestic species); (iv) establishing sample biobanks, making data and samples available for public health risk assessments and mitigation as well as

more detailed pathogen studies; and (v) identifying intervention strategies for human behaviours that increase the risk of novel viral spillover.

## Enhancing our understanding

The Global Virome Project will catalyse new approaches to identify the viruses that represent the greatest threat to human or animal health. The project will use artificial intelligence across the largest viral data set ever assembled, similar to machine learning techniques that are used in genomics to identify gene function, expression and disease biomarkers.<sup>17</sup> The project will use a risk assessment framework that includes data on viral phylogeny, host traits and ecological conditions where the virus exists, as well as a series of viral characteristics known to be associated with spillover, to triage viruses for further characterization.<sup>4</sup> The scale of the project's viral testing will also enable piloting and enhancing novel testing platforms technologies, such as viral capture and sequencing.<sup>18</sup>

The project is ambitious but feasible, enabled by technological developments that allow for rapid and affordable genetic and viral sequencing. The project is time-bound and limited in scope and has tangible progress metrics. In the past, the Human Genome Project, another audacious science project, sequenced and mapped the human genome which focused on the genes with the greatest relevance to people. The ultimate success of the Human Genome Project is in the medical advances made after the project's conclusion. Similarly, the Global Virome Project aims to focus testing on the minimum number of mammalian and waterfowl samples that have the greatest potential of harbouring viruses with zoonotic potential. The exclusion of invertebrates, plants, fish and other hosts from the project's core focus is a deliberate intent to address zoonotic disease. The legacy of the Global Virome Project will probably consist of the countermeasures, diagnostics, vaccines, policies and systems that it enables. Thus, this project has the potential to achieve for pandemics and large-scale epidemics what the Human Genome Project is just beginning to do for personalized medicine.<sup>19</sup>

## Novel countermeasures

The development of countermeasures to viral threats requires significant time and investment, and it is unlikely that these would be developed during the initial Global Virome Project phases. The recently launched Coalition for Epidemic Preparedness Innovations represents a critical step to address known viral threats, such as the Middle East respiratory syndrome coronavirus, Lassa

fever and Nipah virus, for which vaccine or countermeasure development is challenging.<sup>20</sup> The Virome project aims to complement the coalitions' innovations by characterizing the size, structure and composition of the pool of unknown viruses related to the viral targets on which the coalition is focused. For example, if a candidate vaccine against Middle East respiratory syndrome could be tested against hundreds of near relatives of that same syndrome, vaccines that have broader prevention capacity could be selected and rolled out to provide better protection against future emergence of this syndrome. This approach could enhance biotechnological efforts to move from single-virus countermeasures to ones that target a whole family of viruses.<sup>21</sup>

## Project investment

In a single outbreak in one year, the Severe Acute Respiratory Syndrome virus wiped between 10 and 50 billion United States dollars (US \$) of value from Asian stock markets due to disrupted trade and commerce.<sup>22,23</sup> Pandemics are estimated to cause an average of US \$ 570 billion in economic damages per year to the global economy,<sup>24</sup> and these costs will rise as our economies expand and become more interconnected. The Global Virome Project will cost US \$ 1.2 billion, which is less than 0.2% of this estimated loss<sup>4</sup> and less than the estimated US \$ 2.2 billion loss in gross domestic product due to forgone economic growth in Guinea, Liberia, Sierra Leone in 2015 alone during the 2013–2016 Ebola virus disease outbreak.<sup>25</sup> Much like the Global Fund, Gavi, the Vaccine Alliance and other international public health initiatives, the Global Virome Project will likely rely on a broad mix of funding streams from governments, development agencies, research agencies, private foundations and industries. Considering the increasing inevitability of pandemics and their substantial economic impact, the next generation of scientists and field workers trained through this project will have the capacity to monitor viral evolution throughout the coming years. Furthermore, the project's open database will catalyse technological advances in risk assessment, diagnostics and countermeasures.

## Competing interests:

None declared.

## References

1. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, et al. Global trends in emerging infectious diseases. *Nature*. 2008 Feb 21;451(7181):990–3. <https://doi.org/10.1038/nature06536> PMID:18288193
2. The control of neglected zoonotic diseases: a route to poverty alleviation. Geneva: World Health Organization; 2005.
3. Morse SS, Mazet JA, Woolhouse M, Parrish CR, Carroll D, Karesh WB, et al. Prediction and prevention of the next pandemic zoonosis. *Lancet*. 2012 Dec 1;380(9857):1956–65. [http://dx.doi.org/10.1016/S0140-6736\(12\)61684-5](http://dx.doi.org/10.1016/S0140-6736(12)61684-5) PMID:23200504
4. Carroll D, Daszak P, Wolfe ND, Gao GF, Morel CM, Morzaria S, et al. The Global Virome Project. *Science*. 2018 Feb 23;359(6378):872–4. <http://dx.doi.org/10.1126/science.aap7463> PMID:29472471
5. Woolhouse MEJ, Gowtage-Sequeria S. Host range and emerging and reemerging pathogens. *Emerg Infect Dis*. 2005 Dec;11(12):1842–7. <https://doi.org/10.3201/eid1112.050997> PMID:16485468
6. Leroy EM, Kumulungui B, Pourrut X, Rouquet P, Hassanin A, Yaba P, et al. Fruit bats as reservoirs of Ebola virus. *Nature*. 2005 Dec 1;438(7068):575–6. <https://doi.org/10.1038/438575a> PMID:16319873
7. Olival KJ, Hosseini PR, Zambrana-Torrel C, Ross N, Bogich TL, Daszak P. Host and viral traits predict zoonotic spillover from mammals. *Nature*. 2017 06 29;546(7660):646–50. <https://doi.org/10.1038/nature22975> PMID:28636590
8. PREDICT [internet]. Davis: The Regents of the University of California, Davis campus; 2017. Available from: <http://www.predict.global> [cited 2017 Oct 15].
9. PREDICT Consortium. Reducing pandemic risk, promoting global health. Davis: The Regents of the University of California, Davis campus; 2014. Available at: [http://www.vetmed.ucdavis.edu/ohi/local\\_resources/pdfs/predict-global-flyer.pdf](http://www.vetmed.ucdavis.edu/ohi/local_resources/pdfs/predict-global-flyer.pdf) [cited 2017 Oct 9].
10. Olival KJ, Hosseini PR, Zambrana-Torrel C, Ross N, Bogich TL, Daszak P. Host and viral traits predict zoonotic spillover from mammals. *Nature*. 2017 06 29;546(7660):646–50. <https://doi.org/10.1038/nature22975> PMID:28636590
11. Allen T, Murray KA, Zambrana-Torrel C, Morse SS, Rondinini C, Di Marco M, et al. Global hotspots and correlates of emerging zoonotic diseases. *Nat Commun*. 2017 10 24;8(1):1124. <https://doi.org/10.1038/s41467-017-00923-8> PMID:29066781
12. Loh EH, Zambrana-Torrel C, Olival KJ, Bogich TL, Johnson CK, Mazet JA, et al. Targeting transmission pathways for emerging zoonotic disease surveillance and control. *Vector Borne Zoonotic Dis*. 2015 Jul;15(7):432–7. <https://doi.org/10.1089/vbz.2013.1563> PMID:26186515
13. Kreuder Johnson C, Hitchens PL, Smiley Evans T, Goldstein T, Thomas K, Clements A, et al. Spillover and pandemic properties of zoonotic viruses with

14. Ge XY, Li JL, Yang XL, Chmura AA, Zhu G, Epstein JH, et al. Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature*. 2013 Nov 28;503(7477):535–8. <https://doi.org/10.1038/nature12711> PMID:24172901
15. Bird BH, Mazet JAK. Detection of emerging zoonotic pathogens: an integrated One Health Approach. *Annu Rev Anim Biosci*. 2018 Feb 15;6(1):121–39. <http://dx.doi.org/10.1146/annurev-animal-030117-014628> PMID:29144769
16. Chua KB, Gubler DJ. Perspectives of public health laboratories in emerging infectious diseases. *Emerg Microbes Infect*. 2013 Jun;2(6):e37. <https://doi.org/10.1038/emi.2013.34> PMID:26038473
17. Libbrecht MW, Noble WS. Machine learning applications in genetics and genomics. *Nat Rev Genet*. 2015 Jun;16(6):321–32. <https://doi.org/10.1038/nrg3920> PMID:25948244
18. Briese T, Kapoor A, Mishra N, Jain K, Kumar A, Jabado OJ, et al. Virome capture sequencing enables sensitive viral diagnosis and comprehensive virome analysis. *MBiol*. 2015 09 22;6(5):e01491–15. <https://doi.org/10.1128/mBio.01491-15> PMID:26396248
19. Wilson BJ, Nicholls SG. The Human Genome Project, and recent advances in personalized genomics. *Risk Manag Healthc Policy*. 2015 02 16;8:9–20. <https://doi.org/10.2147/RMHP.S58728> PMID:25733939
20. Brede B, Farrar J, Gashumba D, Moedas C, Mundel T, Shiozaki Y, et al. CEPI—a new global R&D organisation for epidemic preparedness and response. *Lancet*. 2017 01 21;389(10066):233–5. [https://doi.org/10.1016/S0140-6736\(17\)30131-9](https://doi.org/10.1016/S0140-6736(17)30131-9) PMID:28109539
21. Warren TK, Wells J, Panchal RG, Stuthman KS, Garza NL, Van Tongeren SA, et al. Protection against filovirus diseases by a novel broad-spectrum nucleoside analogue BCX4430. *Nature*. 2014 Apr 17;508(7496):402–5. <https://doi.org/10.1038/nature13027> PMID:24590073
22. Institute of Medicine Forum on Microbial Threats. Learning from SARS: preparing for the next disease outbreak: workshop summary. Washington (DC): Institute of Medicine; 2004.
23. Flu outbreaks reminder of pandemic threat. Washington (DC): The World Bank; 2013. Available from: <http://www.worldbank.org/en/news/feature/2013/03/05/flu-outbreaks-reminder-of-pandemic-threat> [cited 2017 Oct 14].
24. Fan VY, Jamison DT, Summers LH. The inclusive cost of pandemic influenza risk. NBER working paper no. 22137. Cambridge: National Bureau of Economic Research; 2015. Available from: <http://www.nber.org/papers/w22137> [cited 2017 Oct 10]
25. World Bank Group Ebola response fact sheet. Washington (DC): The World Bank; 2016. Available from: <http://www.worldbank.org/en/topic/health/brief/world-bank-group-ebola-fact-sheet> [cited 2018 Feb 27].