Intercepting pandemics through genomics

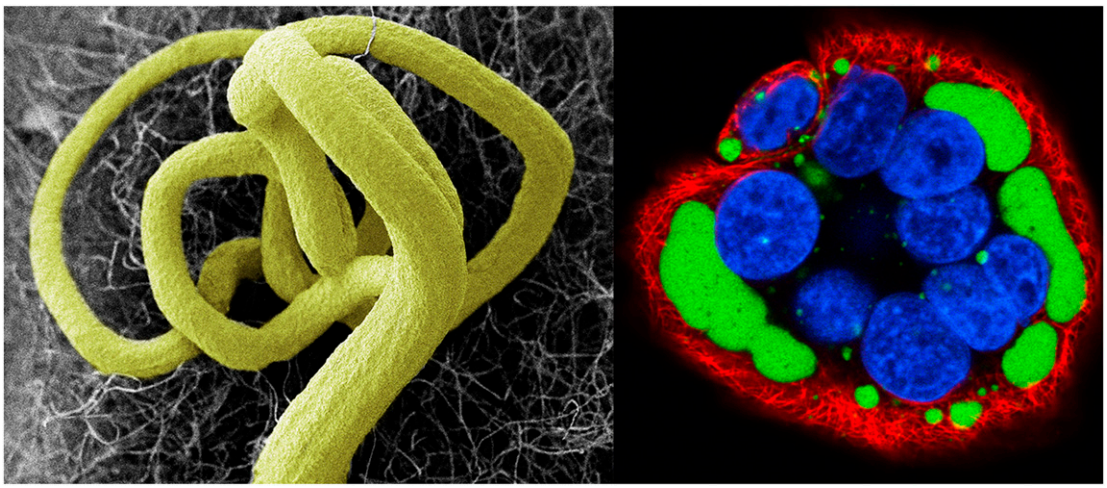
W. John Kress[a,1], Jonna A. K. Mazet[b], and Paul D. N. Hebert[c]

Ecological interactions that cross domains of life have major impacts on ecosystems and human health. Although the coronavirus disease 2019 (COVID-19) pandemic makes this point with destructive clarity, it is clear that zoonotic pathogens pose a standing threat to our species as demonstrated by Ebola, Middle East respiratory syndrome (MERS), and severe acute respiratory syndrome (SARS). Other species experience similar pandemics and are both sources of, and sensitive to, shared pathogens.

Hence, there is an urgent need to establish a global, genomic-based biosurveillance platform, a development which would be of immense value to biosecurity, biodefense, and the economy. If implemented, this "pandemic interception system" would hugely advance our understanding of the natural world.

Three major research programs are poised to support this effort: BIOSCAN, the Earth BioGenome Project (EBP), and the Global Virome Project (GVP). Each of these global programs is now working to develop approaches in comparative genomics that are needed to discover all species and to reveal their interactions. The diversity of infectious agents involved in host–pathogen interactions needs immediate clarification, especially with regard to those agents that transfect phylogenetically divergent lineages. Such information will enable a surveillance system that facilitates preemptive strikes and rapid responses to outbreaks as well as early development of diagnostic pipelines and vaccines. If pursued with a tiny fraction of the resources devoted toward the suppression of COVID-19, rapid progress could be made in identifying every pathogen hosted by birds and mammals with the potential to transfer to humans.

If protecting human life were the sole concern, this effort might be enough. However, our well-being as a species demands environmental sustainability, which can only be achieved by tracking all species as part of our planetary life support system. Achieving this goal means counting species and tracking shifts in their



More than a century ago, *Cryphonectria parasitica* (*Left*), the fungus that causes Chestnut blight, devastated the American Chestnut tree. Until its eradication in 2011, *Rinderpest morbillivirus* (*Right*) decimated cattle and other ungulate populations. A biosurveillance system would help prevent future deadly and economically devastating pathogen outbreaks. Image credit: Paul Beales/UK Crown Copyright (*Left*); © Bioimaging at The Pirbright Institute (*Right*).

[a]National Museum of Natural History, Smithsonian Institution, Washington, DC 20013-7012; [b]One Health Institute, School of Veterinary Medicine, University of California, Davis, CA 95616; and [c]Centre for Biodiversity Genomics, University of Guelph, Guelph, Ontario N1G 2W1, Canada

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: kressj@si.edu.

abundances as well as understanding the role of biotic interactions in driving such changes.

A pandemic interception system needs to be based on detailed knowledge of symbiomes (1), which are the constellations of organisms that interact with all multicellular species. Efforts to describe the structure of symbiomes are motivated by the fact that parasites, parasitoids, and microbes can devastate host populations, especially those that are evolutionarily naïve. Symbiome complexity is governed by rules. Just as large continents support more species than small islands, large-bodied, abundant species have more diverse symbiomes than small, rare taxa.

Now is the time to use the full power of science through cooperative efforts among initiatives such as BIOSCAN, GVP, and EBP to advance our understanding of the complex web of interactions that span the domains of life.

Range expansion also influences symbiome complexity; invasive species soon attract local parasites and microbes. Similarly, anthropogenic change can be a major driver for the emergence of pathogenicity, when host shifts occur as a result of changing land use and the associated encroachment of humans and their domestic animals into previously pristine areas (2). Reflecting our large body size, abundance, and broad distribution, the human symbiome is very diverse. Among our 423 species of eukaryote parasites, just 10% are specialists; most are shared with other mammals or birds (3). The same patterns of cross-taxon associations are true for pathogenic bacteria and viruses (4).

Because of the shifting distributions and abundances of disease-causing agents, symbiomes are in flux. More than 300 emerging disease events have had an impact on human populations over the past 50 years (5). About 44% involved RNA viruses and 25% involved bacteria; three eukaryote lineages (protozoans, fungi, and helminths) accounted for the rest. When viewed from the perspective of pandemic risk, viruses and bacteria are preeminent because they can be transmitted by close contact and aerosols. By contrast, most parasitic eukaryotes are vectored through food (6) or via biting arthropods (7), which are transmission paths that generally limit pandemic potential.

Broadly speaking, two strategies exist for preventing a global pandemic: Either we curb the spread of viral infections once they have started, or we proactively prevent infections by understanding the causes and processes behind pathogen transfers. The latter strategy is more desirable medically, economically, socially, and scientifically. Both actions may be necessary into the foreseeable future.

Indeed, secure systems never rely on a single safeguard, and pandemic protection will require both systematic societal change and scientific progress. Two obvious implementable actions are: 1) halting the trafficking and farming of wildlife to disrupt and prevent the transfer of many pathogens to humans; and 2) temporary cessation of all travel at the first sign of an outbreak to impede the expansion of an epidemic beyond its source region. The public health systems of developed countries should also be reoriented so that they go beyond known pathogens and those affecting their own country. Otherwise diseases, such as COVID-19, can slip through the cracks.

An alternate intervention, one likely to receive broader societal support and to require far less funding than an outbreak response, is the development of a pandemic interception system. Such a system would be founded on comprehensive knowledge of pathogens, their hosts, and their environmental interactions. This knowledge would be coupled with advanced diagnostic methods, and, where justified, anticipatory development of defenses based on detailed evaluation of the susceptibilities of each pathogen.

Part and parcel of a pandemic interception system would be to immediately expand our understanding of the diversity of viral and microbial communities to identify potential pathogens and to clarify their life histories and host interactions before they spill over into vulnerable and susceptible populations (8). To diminish risks to our species, initial work should screen populations of all 15,000 bird and mammal species, because they are the primary hosts for the pathogens most likely to infect humans. A decade ago a project with this scope was impossible, but the exponential rise in output of DNA sequencing platforms has broken this technical barrier.

This research will require a grand alliance. Biodiversity researchers will need to lead specimen acquisition; population biologists will need to track ecological interactions in the field; pathobiologists will need to optimize sampling protocols; genomicists will need to generate complete genomic libraries for all species and screen samples for pathogens. To be fully effective, epidemiologists, computer scientists, and mathematicians will then need to assemble these data to evaluate risks and propose mitigation measures for implementation by public health managers and policymakers. The library of potentially pathogenic agents and their risk characterizations would answer current uncertainties and enable rapid responses. For example, a registry of *Betacoronavirus* lineages would have meant that the primary host of COVID-19 was known long before the current outbreak and would have ensured that diagnostic tests, and perhaps even a vaccine, were rapidly available for deployment.

The foundation for this pandemic interception system is already in place with three ongoing research programs. Since its activation in 2010, the International Barcode of Life (iBOL) Consortium has been using targeted amplicon sequencing to advance the understanding of eukaryote diversity (9). Its current project, BIOSCAN, is radically accelerating species discovery and revealing species interactions by using automated protocols based on minimal sequence

information as well as using multi-gene scans. Involving research organizations with strong capabilities in biodiversity science in more than 30 nations, BIOSCAN is perfectly positioned to collect the required specimens and to verify their identification through DNA-based methods.

Advances in sequencing technologies over the last decade that enable the rapid, efficient, and inexpensive generation of complete genomes led to the emergence of the EBP, which aims to sequence, catalog, and characterize the genomes of all known eukaryotes (10). Currently sequenced genomes are available for fewer than 0.2% of these species. This project will produce new knowledge on the organization, evolution, functions, and interactions among species across the planet. Now a consortium of 32 institutions in 16 countries, EBP is working to generate and analyze genomes across the eukaryotic tree of life to address major issues facing the planet, including the impact of climate change, the conservation of biodiversity, and major perturbations to ecosystem functioning. EBP is very well positioned to provide the genomic data needed to clarify the origins of host–pathogen interactions and the spread of diseases across ecosystems.

The GVP was conceived in 2016 by researchers and policymakers from different disciplines and professional sectors around the globe in response to the repeated, unpredictable emergence and reemergence of high-impact viral epidemics and pandemics compromising global health security and human and animal well-being (11). The GVP benefits from a leadership team with decades of experience in developing and implementing innovative solutions in pandemic prevention and advocating for, and working in, global emerging infectious disease research, policy, and capacity strengthening. The GVP's standard operating procedures and technological and modeling innovations have already enabled the discovery, detection, and risk characterization of 1,200 potentially zoonotic viruses from 35 countries, including more than 100 novel coronaviruses. Its mission will be achieved through a collaborative partnership among public, private, philanthropic, and civil organizations to detect the majority of our planet's unknown virus threats to prepare for and stop future epidemics.

Progress toward a pandemic interception system will be hugely enabled and reinforced by a closer relationship between these three programs. BIOSCAN can lead and support the assembly of properly identified specimens required for analyses and can help to negotiate the complexities linked to sample acquisition. GVP can lead the optimization of both sampling and analytical protocols required to maximize the recovery of novel viruses. EBP can lead and standardize the genome sequence practices needed to better understand interactions within the symbiome structure, the phylogenetic affinities of taxa, and the genomic basis for host susceptibility to pathogens (12). Other global initiatives may well join this effort.

Given the societal and economic disruptions caused by COVID-19, a near-term focus on pathogen surveys related to human health is justified. However, our species is not alone in confronting pandemics. After its introduction into Africa, the rinderpest virus killed millions of cattle, provoking mass starvation in human populations while also decimating all 72 species of African antelopes, which led to severe ecosystem alterations (13). Whereas RNA viruses have often played a central role, fungal pathogens have driven amphibians (14), bats (15), and diverse species of trees (16) to near extinction.

Although the record of past pandemics is clear for larger life, cryptic pandemics likely rage in other groups. Is it possible, for example, that the unexplained recent collapse of global insect populations reflect pandemics linked to introduced pathogens, such as that now decimating honeybee colonies (17)? Pathogenic impacts are well documented for alien parasites that coinvade with their host, but past work has largely focused on eukaryotes. Just as diseases introduced by Europeans decimated the indigenous peoples of the New World (18), invasive species may have unleashed pandemics impacting diverse domains of life across the planet.

Despite the passage of two millennia from the earliest recorded human pandemics (bubonic plague, smallpox, cholera) to their emerging counterparts (H1N1, influenza, SARS, MERS, Ebola, COVID-19), we still react rather than prepare. Currently no coordinated global pandemic interception system exists. Now is the time to use the full power of science through cooperative efforts among initiatives such as BIOSCAN, GVP, and EBP to advance our understanding of the complex web of interactions that span the domains of life. Although most of these interactions are benign or beneficial, a few bring devastation. And while the prevention of future pandemics motivates such efforts, this study of life is sure to bring benefits that extend far beyond the protection of our species to the well-being of our planetary ecosystem.

1 E. A. Tripp *et al.*, Reshaping Darwin's tree: Impact of the symbiome. *Trends Ecol. Evol. (Amst.)* **32**, 552–555 (2017).
2 T. Allen *et al.*, Global hotspots and correlates of emerging zoonotic diseases. *Nat. Commun.* **8**, 1124 (2017).
3 R. W. Ashford, W. Crewe, *The Parasites of Homo Sapiens: An Annotated Checklist of the Protozoa, Helminths and Arthropods for which we are Home* (CRC Press, Liverpool, UK, 1998).
4 M. E. Woolhouse, L. H. Taylor, D. T. Haydon, Population biology of multihost pathogens. *Science* **292**, 1109–1112 (2001).
5 K. E. Jones *et al.*, Global trends in emerging infectious diseases. *Nature* **451**, 990–993 (2008).
6 L. J. Robertson, Parasites in food: From a neglected position to an emerging issue. *Adv. Food Nutr. Res.* **86**, 71–113 (2018).
7 K. F. Smith *et al.*, Global rise in human infectious disease outbreaks. *J. R. Soc. Interface* **11**, 20140950 (2014).
8 C. K. Johnson *et al.*, Global shifts in mammalian population trends reveal key predictors of virus spillover risk. *Proc. Biol. Sci.* **287**, 20192736 (2020).
9 P. D. N. Hebert, P. M. Hollingsworth, M. Hajibabaei, From writing to reading the encyclopedia of life. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **371**, 20150321 (2016).

10 H. A. Lewin *et al.*, Earth BioGenome Project: Sequencing life for the future of life. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 4325–4333 (2018).
11 D. Carroll *et al.*, The Global Virome Project. *Science* **359**, 872–874 (2018).
12 J. Damas *et al.*, Broad host range of SARS-CoV-2 predicted by comparative and structural analysis of ACE2 in vertebrates. bioRxiv: 10.1101/2020.04.16.045302 (18 April 2020).
13 R. M. Holdo *et al.*, A disease-mediated trophic cascade in the Serengeti and its implications for ecosystem C. *PLoS Biol.* **7**, e1000210 (2009).
14 A. McDermott, News Feature: Fighting a fungal scourge. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 20245–20249 (2019).
15 J. Zukal *et al.*, White-nose syndrome fungus: a generalist pathogen of hibernating bats. *PLoS One* **9**, e97224 (2014).
16 A. Santini, A. Battisti, Complex insect-pathogen interactions in tree pandemics. *Front. Physiol.* **10**, 550 (2019).
17 G. E. Budge *et al.*, Chronic bee paralysis as a serious emerging threat to honey bees. *Nat. Commun.* **11**, 2164 (2020).
18 R. S. Walker, L. Sattenspiel, K. R. Hill, Mortality from contact-related epidemics among indigenous populations in Greater Amazonia. *Sci. Rep.* **5**, 14032 (2015).