international bodies promoting standards in microbiome research.

A number of large data-management systems are currently available for supporting the comparative analysis of assembled [12] or unassembled [13] microbiome data and their associated metadata [14], as well as systems designed for predictive modeling (https://kbase.us/) and cyberinfrastructures [15]. Similar successful systems with existing and dedicated long-term funding should be an integral part of such a distributed national microbiome data center.

Concluding Remarks

Future endeavors in microbiome research are expected to lead us to a new age of holistic understanding of microbial life, develop novel therapeutic strategies to treat infectious diseases, identify solutions for protecting the environment, and ultimately understand and harness the power of the most abundant natural resources on our planet. To achieve these endeavors and enable the vision described above, the research community requires a major restructuring in the current research-funding policies through the development of innovative funding mechanisms that will provide long-term support for microbiome data science. Examples of such mechanisms can be drawn from existing models such as the Brain Initiative (https://www.whitehouse.gov/share/brain-initiative), a grand challenge research effort to revolutionize our understanding of the human brain. At the dawn of the third decade of microbial genomics, and well into the information age, the time is ripe to embark on the greatest endeavor to understand Earth's microbiome. Microbiome data science, through the establishment of a national microbiome data center, can pave the way.

Acknowledgments

We thank Victor Markowitz, Torben Nielsen, and Heather Maughan for critical reading and suggestions on the manuscript. This work was conducted by the U. S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, under Contract No. DE-AC02-05CH11231.

[1]Prokaryotic Super Program, Department of Energy Joint Genome Institute, Walnut Creek, CA, USA

*Correspondence: nckyrpides@lbl.gov (N.C. Kyrpides).

http://dx.doi.org/10.1016/j.tim.2016.02.011

References

1. Koboldt, D.C. *et al.* (2013) The next-generation sequencing revolution and its impact on genomics. *Cell* 155, 27–38
2. Kyrpides, N.C. (2009) Fifteen years of microbial genomics: meeting the challenges and fulfilling the dream. *Nat. Biotechnol.* 27, 627–632
3. Alivisatos, A.P. *et al.* (2015) A unified initiative to harness Earth's microbiomes. *Science* 350, 507–508
4. Dubilier, N. *et al.* (2015) Microbiology: Create a global microbiome effort. *Nature* 526, 631–634
5. Wu, D. *et al.* (2009) A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature* 462, 1056–1060
6. Kyrpides, N.C. *et al.* (2014) Genomic encyclopedia of bacteria and archaea: sequencing a myriad of type strains. *PLoS Biol.* 12, e1001920
7. Gilbert, J.A. *et al.* (2014) Meeting report: Ocean 'omics science, technology and cyberinfrastructure: current challenges and future requirements (August 20–23, 2013). *Stand. Genomic Sci.* 9, 1252–1258
8. Field, D. *et al.* (2011) The Genomic Standards Consortium. *PLoS Biol.* 9, e1001088
9. Field, D. *et al.* (2011) Genomic standards consortium projects. *Stand. Genomic Sci.* 9, 514–526
10. National Research Council (US) Committee on Metagenomics: Challenges and Functional Applications (2007) *The New Science of Metagenomics: Revealing the Secrets of Our Microbial Planet,* National Academies Press (US)
11. Ioannidis, J.P.A. (2005) Why most published research findings are false. *PLoS Med.* 2, 696–701
12. Markowitz, V.M. *et al.* (2014) IMG/M 4 version of the integrated metagenome comparative analysis system. *Nucleic Acids Res.* 42, D568–D573
13. Wilke, A. *et al.* (2013) A metagenomics portal for a democratized sequencing world. *Meth. Enzymol.* 531, 487–523
14. Reddy, T.B. *et al.* (2015) The Genomes OnLine Database (GOLD) v.5: a metadata management system based on a four level (meta)genome project classification. *Nucleic Acids Res.* 43, D1099–D1106
15. Goff, S.A. *et al.* (2011) The iPlant Collaborative: cyberinfrastructure for plant biology. *Front. Plant Sci.* 2, 34

Spotlight

Engineering Coronaviruses to Evaluate Emergence and Pathogenic Potential

Susanna K.P. Lau[1,2,3,4,5,*] and Patrick C.Y. Woo[1,2,3,4,5,*]

A recent study provides a platform for generating infectious coronavirus genomes using sequence data, examining their capabilities of replicating in human cells and causing diseases in animal models, and evaluating therapeutics and vaccines. Similar approaches could be used to assess the potential of human emergence and pathogenicity for other viruses.

The severe acute respiratory syndrome (SARS) epidemic in 2003 and the Middle East respiratory syndrome (MERS) epidemic in the last 3 years have shown that coronaviruses (CoVs) have the capability to cause major epidemics. For the SARS epidemic, a total of >8000 laboratory-confirmed cases with >800 deaths were observed (http://www.cdc.gov/sars/about/fs-sars.html). This horrific epidemic was followed by the publication of >7500 scientific papers on CoVs visible in PubMed, which represents two-thirds of the total number of publications on CoVs in Pubmed. Despite the numerous studies on CoVs, it is still difficult to predict which CoV may have the potential to emerge as the next culprit. A recent study in *PNAS* by Menachery *et al.* [1] and another similar study in *Nature Medicine* published in December 2015 by the same group [2] reported the use of existing sequence data with reverse genetics to engineer SARS-related CoVs and evaluate their potential of emergence and pathogenicity.

Shortly after the emergence of SARS-CoV, SARS-related CoVs were found in civets [3]. However, multiple lines of evidence showed that the civets are just the intermediate or amplification hosts for SARS-CoV. Through intensive surveillance studies in various mammals in Hong Kong, Lau *et al.* reported the presence of SARS-related CoVs in Chinese horseshoe bats in Hong Kong [4]. A similar observation was also reported by another group in mainland China [5]. Since then, numerous SARS-related CoV sequences were observed in different

species of the horseshoe bat. These SARS-related CoV sequences possessed different degrees of sequence identities to the SARS-CoV originally found in humans. Most importantly, it was difficult to predict which SARS-related CoV may have the potential to emerge in humans again, causing another SARS epidemic.

In order to predict whether a SARS-related bat CoV, named WIV1-CoV, discovered in Chinese horseshoe bats in Yunnan [6], had the potential to emerge in humans, Menachery *et al.* synthesized a full-DNA clone based on the sequence of the virus [1]. The clone was shown to be able to generate infectious virus in primary human airway epithelial cell culture, confirming its ability to efficiently replicate in human cells. When the spike gene of WIV1-CoV was used to replace that of SARS-CoV MA15, a mouse-adapted virus [7], the resulting CoV did not replicate efficiently or cause disease in mice. However, if transgenic mice that expressed the ACE2 receptor were used, WIV1-CoV was able to replicate and cause weight loss. These experiments showed that WIV1-CoV actually has significant potential to emerge in humans.

In the second part of their study, Menachery *et al.* evaluated various therapeutics and vaccines for possible treatment and prevention of WIV1-CoV infection [1]. They found that antibodies that have been shown to block SARS-CoV might also be used for protection of humans against WIV1-CoV infections. However, they found that the antibodies generated by immunizing animals using inactivated SARS-CoV did not protect WIV1-CoV infection in aged mice. This implied that, if WIV1-CoV emerges to cause another epidemic, immunization of humans using inactivated SARS-CoV may not be useful to control the epidemic.

The platform that Menachery *et al.* employed can be used for evaluating the emergence and pathogenic potential of other CoVs. Before the SARS epidemic, fewer than 10 CoVs with complete genome sequences were available. After the SARS epidemic, and up to March 2016, there was an addition of more than 40 CoVs with complete genomes sequenced. These CoVs include two human CoVs (HCoVs; HCoV-NL63 and HCoV-HKU1) and at least 30 other mammalian and avian CoVs. Two additional lineages in *Betacoronavirus* and a novel genus, *Deltacoronavirus*, have been discovered [8,9]. This diversity of coronaviruses is due to the infidelity of their RNA-dependent RNA polymerases, their high frequencies of homologous RNA recombination, and their large genomes. CoVs are well known for the difficulties in culturing them, and most of the newly discovered CoVs are so far noncultivable. Just from the sequences of these CoVs, it is impossible to predict which one(s) may have the potential in causing the next epidemic. Synthetic full-DNA clones for these CoVs can be made, and systematic evaluation can be performed. Since betaCoVs, including HCoV-OC43, SARS-CoV, and MERS-CoV, are the culprits of all large human epidemics, we should start the systematic evaluation by focusing on the betaCoVs.

In addition to CoVs, this approach can also be extended to study other viruses with the potential for emergence. In recent years, multiple viruses – such as influenza viruses H5N1 and H7N9, SARS-CoV and MERS-CoV, Ebolavirus and Zika virus – have emerged in humans. Some of them have apparently disappeared after the amplification host was segregated from humans, whereas others have persisted for years. Each of these viruses has a lot of closely related viruses or sequences obtained by either conventional sequencing or metagenomics approaches. The platform of Menachery *et al.* can be used for assessing the potential of these cultivable viruses or viral sequences in causing another epidemic, and the effectiveness of using various therapeutic or vaccination modalities in their treatment or prevention if an epidemic due to these viruses really emerges.

Despite the academic contribution of these research studies, such potentially gain-of-function experiments must be scrutinized and handled with the utmost care. Recently, there has been much discussion and debate on the potential threats of gain-of-function experiments [10]. There are fears that accidental release of these 'super-virulent' viruses or their use as biological weapons may lead to uncontrollable epidemics. However, we believe that these gain-of-function experiments should not be banned because of the potential threats. By contrast, each experiment should be carefully examined for the potential benefits weighed against potential threats. In our opinion, the potential benefits are enormous, as exemplified by the works by Menachery *et al.* [1,2]. It is the laboratory practice and conduct of the researchers that are most important in safeguarding any potential leakage of viruses.

Acknowledgments

This work is partly supported by the Theme-based Research Scheme (Project No. T11/707/15), University Grant Committee; and University Development Fund and Strategic Research Theme Fund, The University of Hong Kong. We thank Hayes Luk for the useful discussions.

[1]State Key Laboratory of Emerging Infectious Diseases, The University of Hong Kong, Hong Kong, China
[2]Department of Microbiology, The University of Hong Kong, Hong Kong, China
[3]Research Centre of Infection and Immunology, The University of Hong Kong, Hong Kong, China
[4]Carol Yu Centre for Infection, The University of Hong Kong, Hong Kong, China
[5]Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Zhejiang University, Hangzhou 310006, China

*Correspondence: skplau@hku.hk (Susanna K.P. Lau) and pcywoo@hku.hk (Patrick C.Y. Woo).

http://dx.doi.org/10.1016/j.tim.2016.04.001

References

1. Menachery, V.D. *et al.* (2016) SARS-like WIV1-CoV poised for human emergence. *Proc. Natl. Acad. Sci. U.S.A.* 113, 3048–3053
2. Menachery, V.D. *et al.* (2015) A SARS-like cluster of circulating bat coronaviruses shows potential for human emergence. *Nat. Med.* 21, 1508–1513
3. Guan, Y. *et al.* (2003) Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. *Science* 302, 276–278
4. Lau, S.K. *et al.* (2005) Severe acute respiratory syndrome coronavirus-like virus in Chinese horseshoe bats. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14040–14045
5. Li, W. *et al.* (2005) Bats are natural reservoirs of SARS-like coronaviruses. *Science* 310, 676–679

6. Ge, X.Y. *et al.* (2013) Isolation and characterization of a bat SARS-like coronavirus that uses the ACE2 receptor. *Nature* 503, 535–538
7. Roberts, A. *et al.* (2007) A mouse-adapted SARS-coronavirus causes disease and mortality in BALB/c mice. *PLoS Pathog.* 3, e5
8. Woo, P.C. *et al.* (2007) Comparative analysis of twelve genomes of three novel group 2c and group 2d coronaviruses reveals unique group and subgroup features. *J. Virol.* 81, 1574–1585
9. Woo, P.C. *et al.* (2012) Discovery of seven novel Mammalian and avian coronaviruses in the genus *Deltacoronavirus* supports bat coronaviruses as the gene source of *Alphacoronavirus* and *Betacoronavirus* and avian coronaviruses as the gene source of *Gammacoronavirus* and *Deltacoronavirus*. *J. Virol.* 86, 3995–4008
10. Kilianski, A. *et al.* (2016) Gain-of-function research and the relevance to clinical practice. *J. Infect. Dis.* 213, 1364–1369

Spotlight

West Nile Virus Fitness Costs in Different Mosquito Species

Lark L. Coffey[1,*] and William K. Reisen[1]

West Nile virus (WNV) remains an important public health problem causing annual epidemics in the United States. Grubaugh *et al*. observed that WNV genetic divergence is dependent on the vector mosquito species. This suggests that specific WNV vector–bird species pairings may generate novel genotypes that could promote outbreaks.

Ongoing outbreaks of arthropod-borne (arbo) viruses highlight the need for understanding their transmission and emergence. Zika (ZIKV), chikungunya (CHIKV), and West Nile (WNV) viruses have recently emerged in new regions of the world to cause significant human disease. Despite garnering less media attention recently than ZIKV and CHIKV, WNV remains an important public health threat to the USA, and, unlike ZIKV and CHIKV, has established local transmission and continues to cause epidemics across the continental USA every summer since it was introduced in 1999. In fact, the ongoing WNV epidemic has resulted in the largest domestic arboviral neuroinvasive disease outbreak on record, causing 16000 cases with more than 1600 deaths [1]. However, in contrast to ZIKV and CHIKV that use humans as amplifying hosts and two primary urban mosquito vectors, WNV uses several avian species as amplifying hosts and at least four major vector mosquito species [1]. The ability of WNV to infect many species suggests this virus is especially capable of surviving unique evolutionary pressures compared to more host-restricted arboviruses that use fewer host species. Despite this, the effects of catholic host specificity on WNV evolution and emergence have been poorly studied to date.

Successful viral transmission by vector mosquitoes necessitates viral replication within mosquito tissues, dissemination through a variety of organs, and expectoration in saliva during subsequent bloodfeeding. Infection is dose dependent such that mosquitoes ingest many virus particles but transmit few during re-feeding. This reduction in viral population size produces genetic bottlenecks and founder effects that present a challenge to survival and influence viral evolution.

Almost all arboviruses have RNA genomes whose viral polymerases are unable to error-correct, resulting in about one mutation per genome replicated [2]. This mutability, coupled with rapid evolution, helps RNA viruses ensure their survival. Mutated genomes together form heterogenous intrahost populations that are highly similar but not identical. Although the majority of mutations probably hurt the virus, being mutable also allows arboviruses to adapt to new settings when mutations at minority frequencies in populations are positively selected in new or different environments and rise to dominate the population. In addition to helping the virus survive disparate vertebrate and invertebrate hosts, positively selected mutations can also promote an increase in disease.

Advances in deep sequencing have enabled genome characterization, revealing that intrahost arbovirus populations in humans or birds and mosquitoes are different [3–5]. However, viral population dynamics within mosquito vectors are poorly understood for any given vector, and no studies have compared viral populations across vector species. For WNV, genetic variation could allow the virus to successfully traverse different adaptive landscapes presented by its many bird and mosquito hosts. Recent studies with CHIKV showed that what goes into the mosquito does not always reflect what comes out during re-feeding. Mutations that only dominate populations in saliva may enhance vector transmissibility [6], a factor that could promote further CHIKV spread and contribute to explosive outbreaks. However, similar studies have not been performed for WNV, including using different species that represent the range of primary vectors the virus uses.

Recently in *Cell Host & Microbe*, Grubaugh *et al*. [7] used deep sequencing to characterize WNV populations in tissues from four vector mosquitoes: three enzootic species *Culex tarsalis*, *Cx. quinquefasciatus*, and *Cx. pipiens*, and the potential bridge vector, *Aedes aegypti.* The goal was to assess how WNV populations evolve in different mosquito vectors. The four species were fed the same WNV-spiked bloodmeal and then midguts (representing infection), legs (disseminated infections), salivary glands, and saliva (transmission) from infected mosquitoes were analyzed to study intratissue WNV population structure. Predictably, individuals of all four species with more WNV genomes were more likely to transmit WNV. What was less predictable, and also novel to this study, is that the intrahost population structure was unique to each mosquito species. *Cx. quinquefasciatus* and *Cx. tarsalis* developed more