

Fruits of Virus Discovery: New Pathogens and New Experimental Models

David Wang

Departments of Molecular Microbiology and Pathology & Immunology, Washington University School of Medicine, St. Louis, Missouri, USA

The advent of high-throughput sequencing has led to a tremendous increase in the rate of discovery of viral sequences. In some instances, novel pathogens have been identified. What has been less well appreciated is that novel virus discoveries in distinct hosts have led to the establishment of unique experimental systems to define host-virus interactions. These new systems have opened new frontiers in the study of fundamental virology and infectious disease.

NEXT-GENERATION SEQUENCING (NGS) CATALYZES THE RATE OF VIRUS DISCOVERY

In the past decade, new developments in sequencing technology, along with commensurate bioinformatic capability and tools and their application to interrogate samples for the presence of virus-derived nucleic acids, have led to a dramatic increase in the rate at which novel virus sequences have been discovered. In single studies of diverse specimen types, evidence for literally tens to hundreds of novel viruses can now be generated. This has resulted in a near-quadrupling of the number of unique virus species in GenBank since 2005 (Fig. 1). These observations have demonstrated that the world of viruses is vastly greater than previously recognized. For the past half-century, the predominant viruses that were known and studied were those that could be grown in cell culture systems, and it was only on rare occasions that viruses were identified that could not be cultured (e.g., human norovirus or hepatitis C). Thus, our perspective on the diversity of viruses has been heavily biased toward those that could be propagated in a very limited set of experimental culture systems. In many regards, this situation parallels the 1990-era observations in the bacterial world when consensus 16S rRNA gene sequencing demonstrated a much greater diversity of bacterial taxa than had been previously identified by culturing.

The goal of this Gem is to describe the most substantive fruits arising from the NGS revolution. Technical aspects of viral discovery by NGS are not discussed here. Rather, the focus is on providing exemplars of novel pathogens and novel model systems that have emerged from the first decade of NGS-based virus-discovery efforts.

IDENTIFICATION OF NOVEL PATHOGENS

Many novel viruses have been identified by application of NGS to diseased specimens, thereby generating candidate agents that may play causal roles in the disease in question. Two examples that illustrate new paradigms emerging from virus-discovery efforts are presented.

Merkel polyomavirus and human cancer. For decades, there has been controversy as to whether polyomaviruses, which have well-defined transforming properties in animal models, play causal roles in any human tumors. In 2008, NGS analysis of specimens collected from Merkel cell carcinoma, a rare tumor that develops more frequently in immunocompromised individuals, identified a novel polyomavirus. Merkel polyomavirus was clonally integrated in multiple tumors, and subsequent analyses dem-

onstrated the presence of this virus in ~80% of Merkel cell carcinomas tested (1). The data accumulated to date provide the strongest evidence implicating a polyomavirus in the etiology of a human cancer. Furthermore, detailed analysis of the Merkel genome led to the identification of a novel open reading frame termed ALTO (2). While its function is still cryptic, comparative genomic analyses suggest that there is an entire clade of polyomaviruses that encode ALTO. Thus, from this discovery, not only was a novel pathogen identified, but new insights into the fundamental virologic and evolutionary properties of polyomaviruses that were not previously recognized were gleaned.

Astrovirus VA1 and PS. Astroviruses have traditionally been associated with gastrointestinal disease in humans and animals. Therefore, it was not surprising when application of NGS to stool samples from an unexplained outbreak of gastroenteritis in a Virginia day care center led to the discovery of a novel astrovirus, astrovirus VA1, that was present in multiple samples from the outbreak (3). More surprising was the identification of astrovirus PS, a virus that shares 95% nucleotide identity with astrovirus VA1, in brain tissue of an immunocompromised child with unexplained encephalitis (4). The detection of this novel astrovirus in brain tissue strongly implicates this virus as the cause of the patient's symptoms and represents the first example of astrovirus-associated neurologic disease in humans. Interestingly, additional NGS studies have also identified other novel astroviruses associated with neurologic diseases of cows and mink (5, 6). Taken together, these NGS-based studies have expanded our understanding of the tissue tropism and potential disease spectrum of astroviruses.

ESTABLISHMENT OF NEW ANIMAL EXPERIMENTAL MODELS

Aside from identifying novel pathogens, a key benefit from the flurry of virus discovery activities has been the establishment of new animal model systems for understanding viral pathogenesis

Accepted manuscript posted online 19 November 2014

Citation Wang D. 2015. Fruits of virus discovery: new pathogens and new experimental models. *J Virol* 89:1486–1488. doi:10.1128/JVI.01194-14.

Editor: C. S. Sullivan

Address correspondence to davewang@wustl.edu.

Copyright © 2015, American Society for Microbiology. All Rights Reserved.
doi:10.1128/JVI.01194-14

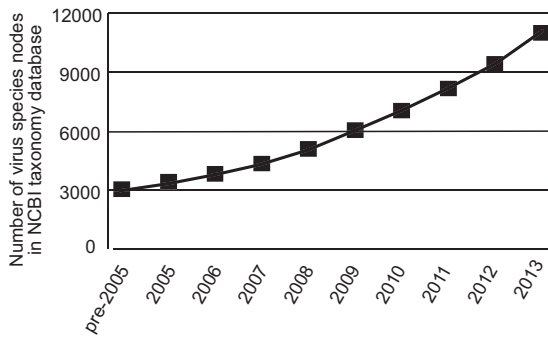


FIG 1 Novel virus sequences deposited in GenBank based upon NCBI taxonomy criteria. The cumulative numbers of unique virus species nodes are plotted.

and virus-host interactions. This includes model systems for specific pathogens and more generalized models to examine broader questions of fundamental host-virus interactions.

Animal hepatitis C models. For many years, studies of hepatitis C were limited by the absence of a cell culture system and the lack of small-animal models. The recent discoveries of rodent (7, 8) and canine (9) homologs of hepatitis C virus provide avenues to develop new small-animal models for the study of these relatives of hepatitis C virus. While these studies are only in their infancy, paradigms for the effectiveness of such parallel rodent models include studies of murine gammaherpesvirus 68 as a surrogate for understanding Epstein-Barr virus and Kaposi's sarcoma herpesvirus and the use of murine norovirus as a model for human norovirus.

Bats as reservoirs of viral diversity. A major finding of the past decade has been the identification of bats as a source of tremendous virus diversity. Catalyzed in large part by the discovery of severe acute respiratory syndrome (SARS)-like coronavirus in bats, subsequent NGS- and taxon-specific consensus PCR-based virus-discovery efforts have demonstrated that bats harbor novel sequences derived from many known viral families (10–12). However, efforts to progress beyond discovery to experimentally characterize these viruses have been hampered by the general inability to cultivate the majority of the novel bat viruses in common mammalian cell lines. Thus, significant effort in recent years has focused on developing bat reagents—including genomics, transcriptomics, and new immortalized bat cell lines and primary bat cultures (13–16). Such tools are critically important for defining the fundamental aspects of viral infection in bats, which in turn will help to resolve issues such as what makes bats such permissive hosts of viral infection and what factors govern cross-species transmission of viruses from bats. It seems likely that with the development of appropriate bat cell culture systems, the experimental tractability of bat viruses will mature much as the development of insect cell cultures in the 1960s revolutionized the culture of arboviruses.

Virus infection in the model organism *C. elegans*. *Caenorhabditis elegans* is a highly genetically tractable model organism that has played key roles in the discoveries of broadly conserved fundamental processes such as the caspase cell death pathways and RNA interference. However, its utility in the study of virus-host interactions has been limited due to the absence of any viruses known to naturally infect *C. elegans*. In fact, although nematodes are among the most abundant and diverse animal species on earth,

there were no viruses naturally infecting any nematode that had been molecularly characterized until 2011, when Orsay virus, the first and, to date, only known virus capable of infecting *C. elegans*, was discovered (17). Orsay virus is a positive-sense RNA virus most similar to nodaviruses. Experimental conditions for infection of the laboratory reference *C. elegans* strain, as well as a viral reverse genetic system for Orsay virus (18), have now been established, providing a robust new model system to define proviral and antiviral factors against this RNA virus. Such studies may not only provide new insights into invertebrate virology; because ~40% of the genes in *C. elegans* are conserved in mammals, the system also has the potential to identify host pathways that are broadly evolutionarily conserved. The discovery of Toll in *Drosophila* and the subsequent characterization of Toll-like receptors in vertebrates provide a paradigm for the power of model organism studies of infection and immunity.

FUTURE DIRECTIONS AND CHALLENGES

It is clear that the advent of NGS has triggered a new age of discovery of novel virus sequences. Continued application to humans, animals, plants, and environmental niches will continue to uncover novel virus diversity for some time. Recent estimates for mammals suggest that on the order of 10^5 novel viruses remain to be discovered (19), which would provide much fodder for future studies. These include traditional efforts to define pathogenic roles of novel viruses in diseases as well as the exploitation of new genomes to identify novel genes, properties, or host interactions absent in their previously described relatives by comparative analyses. Furthermore, additional novel experimental models will undoubtedly arise from the multitude of new viruses identified. One limitation is that current NGS-based virus-discovery methods still rely largely upon alignment to known viruses; thus, truly novel viruses from previously unrecognized taxa will continue to evade detection until robust alignment-independent bioinformatic tools are developed. In addition, the rate of discovery of new viral sequences has rapidly outpaced the rate at which culture systems for viruses have been developed, thereby limiting experimental studies of these new viruses. To maximally benefit from the wealth of viruses being discovered, there must be commensurate effort and resources dedicated to building fundamental tools and reagents for propagating these novel viruses. Despite these challenges, there will clearly be more low-hanging fruit to be picked and exploited to advance our understanding of fundamental virology and viral pathogenesis.

ACKNOWLEDGMENTS

Work in my laboratory relevant to this topic is supported in part by NIH grants U54 AI057160 and R21 AI097865 and by an Investigator in the Pathogenesis of Infectious Diseases Award from the Burroughs Wellcome Fund.

REFERENCES

- Feng H, Shuda M, Chang Y, Moore PS. 2008. Clonal integration of a polyomavirus in human Merkel cell carcinoma. *Science* 319:1096–1100. <http://dx.doi.org/10.1126/science.1152586>.
- Carter JJ, Daugherty MD, Qi X, Bheda-Malge A, Wipf GC, Robinson K, Roman A, Malik HS, Galloway DA. 2013. Identification of an overprinting gene in Merkel cell polyomavirus provides evolutionary insight into the birth of viral genes. *Proc Natl Acad Sci U S A* 110:12744–12749. <http://dx.doi.org/10.1073/pnas.1303526110>.
- Finkbeiner SR, Li Y, Ruone S, Conrardy C, Gregoricus N, Toney D, Virgin HW, Anderson LJ, Vinje J, Wang D, Tong S. 2009. Identification

- of a novel astrovirus (astrovirus VA1) associated with an outbreak of acute gastroenteritis. *J Virol* 83:10836–10839. <http://dx.doi.org/10.1128/JVI.00998-09>.
4. Quan PL, Wagner TA, Briese T, Torgerson TR, Hornig M, Tashmukhamedova A, Firth C, Palacios G, Baisre-De-Leon A, Paddock CD, Hutchison SK, Egholm M, Zaki SR, Goldman JE, Ochs HD, Lipkin WI. 2010. Astrovirus encephalitis in boy with X-linked agammaglobulinemia. *Emerg Infect Dis* 16:918–925. <http://dx.doi.org/10.3201/eid1606.091536>.
 5. Blomström AL, Widen F, Hammer AS, Belak S, Berg M. 2010. Detection of a novel astrovirus in brain tissue of mink suffering from shaking mink syndrome by use of viral metagenomics. *J Clin Microbiol* 48:4392–4396. <http://dx.doi.org/10.1128/JCM.01040-10>.
 6. Li L, Diab S, McGraw S, Barr B, Traslavina R, Higgins R, Talbot T, Blanchard P, Rimoldi G, Fahsbender E, Page B, Phan TG, Wang C, Deng X, Pesavento P, Delwart E. 2013. Divergent astrovirus associated with neurologic disease in cattle. *Emerg Infect Dis* 19:1385–1392. <http://dx.doi.org/10.3201/eid1909.130682>.
 7. Drexler JF, Corman VM, Muller MA, Lukashev AN, Gmyl A, Coutard B, Adam A, Ritz D, Leijten LM, van Riel D, Kallies R, Klose SM, Gloza-Rausch F, Binger T, Annan A, Adu-Sarkodie Y, Oppong S, Bourgairel M, Rupp D, Hoffmann B, Schlegel M, Kummerer BM, Kruger DH, Schmidt-Chanasit J, Setien AA, Cottontail VM, Hemachudha T, Wacharapluesadee S, Osterrieder K, Bartenschlager R, Matthee S, Beer M, Kuiken T, Reusken C, Leroy EM, Ulrich RG, Drosten C. 2013. Evidence for novel hepaciviruses in rodents. *PLoS Pathog* 9:e1003438. <http://dx.doi.org/10.1371/journal.ppat.1003438>.
 8. Kapoor A, Simmonds P, Scheel TK, Hjelle B, Cullen JM, Burbelo PD, Chauhan LV, Duraisamy R, Sanchez Leon M, Jain K, Vandegriff KJ, Calisher CH, Rice CM, Lipkin WI. 2013. Identification of rodent homologs of hepatitis C virus and pegiviruses. *mBio* 4:e00216-13. <http://dx.doi.org/10.1128/mBio.00216-13>.
 9. Kapoor A, Simmonds P, Gerold G, Qaisar N, Jain K, Henriquez JA, Firth C, Hirschberg DL, Rice CM, Shields S, Lipkin WI. 2011. Characterization of a canine homolog of hepatitis C virus. *Proc Natl Acad Sci U S A* 108:11608–11613. <http://dx.doi.org/10.1073/pnas.1101794108>.
 10. Donaldson EF, Haskew AN, Gates JE, Huynh J, Moore CJ, Frieman MB. 2010. Metagenomic analysis of the viromes of three North American bat species: viral diversity among different bat species that share a common habitat. *J Virol* 84:13004–13018. <http://dx.doi.org/10.1128/JVI.01255-10>.
 11. Li L, Victoria JG, Wang C, Jones M, Fellers GM, Kunz TH, Delwart E. 2010. Bat guano virome: predominance of dietary viruses from insects and plants plus novel mammalian viruses. *J Virol* 84:6955–6965. <http://dx.doi.org/10.1128/JVI.00501-10>.
 12. Zhu HC, Chu DK, Liu W, Dong BQ, Zhang SY, Zhang JX, Li LF, Vijaykrishna D, Smith GJ, Chen HL, Poon LL, Peiris JS, Guan Y. 2009. Detection of diverse astroviruses from bats in China. *J Gen Virol* 90:883–887. <http://dx.doi.org/10.1099/vir.0.007732-0>.
 13. Crameri G, Todd S, Grimley S, McEachern JA, Marsh GA, Smith C, Tachedjian M, De Jong C, Virtue ER, Yu M, Bulach D, Liu JP, Michalski WP, Middleton D, Field HE, Wang LF. 2009. Establishment, immortalisation and characterisation of pteropid bat cell lines. *PLoS One* 4:e8266. <http://dx.doi.org/10.1371/journal.pone.0008266>.
 14. Eckerle I, Ehlen L, Kallies R, Wollny R, Corman VM, Cottontail VM, Tschapka M, Oppong S, Drosten C, Muller MA. 2014. Bat airway epithelial cells: a novel tool for the study of zoonotic viruses. *PLoS One* 9:e84679. <http://dx.doi.org/10.1371/journal.pone.0084679>.
 15. Papenfuss AT, Baker ML, Feng ZP, Tachedjian M, Crameri G, Cowled C, Ng J, Janardhana V, Field HE, Wang LF. 2012. The immune gene repertoire of an important viral reservoir, the Australian black flying fox. *BMC Genomics* 13:261. <http://dx.doi.org/10.1186/1471-2164-13-261>.
 16. Zhang G, Cowled C, Shi Z, Huang Z, Bishop-Lilly KA, Fang X, Wynne JW, Xiong Z, Baker ML, Zhao W, Tachedjian M, Zhu Y, Zhou P, Jiang X, Ng J, Yang L, Wu L, Xiao J, Feng Y, Chen Y, Sun X, Zhang Y, Marsh GA, Crameri G, Broder CC, Frey KG, Wang LF, Wang J. 2013. Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. *Science* 339:456–460. <http://dx.doi.org/10.1126/science.1230835>.
 17. Félix MA, Ashe A, Piffaretti J, Wu G, Nuez I, Bécicard T, Jiang Y, Zhao G, Franz CJ, Goldstein LD, Sanroman M, Miska EA, Wang D. 2011. Natural and experimental infection of caenorhabditis nematodes by novel viruses related to nodaviruses. *PLoS Biol* 9:e1000586. <http://dx.doi.org/10.1371/journal.pbio.1000586>.
 18. Jiang H, Franz CJ, Wang D. 2014. Engineering recombinant Orsay virus directly in the metazoan host *Caenorhabditis elegans*. *J Virol* 88:11774–11781. <http://dx.doi.org/10.1128/JVI.01630-14>.
 19. Anthony SJ, Epstein JH, Murray KA, Navarrete-Macias I, Zambrana-Torrel CM, Solovyov A, Ojeda-Flores R, Arrigo NC, Islam A, Ali Khan S, Hosseini P, Bogich TL, Olival KJ, Sanchez-Leon MD, Karesh WB, Goldstein T, Luby SP, Morse SS, Mazet JA, Daszak P, Lipkin WI. 2013. A strategy to estimate unknown viral diversity in mammals. *mBio* 4:e00598-13. <http://dx.doi.org/10.1128/mBio.00598-13>.