JVI Accepts, published online ahead of print on 27 August 2014
J. Virol. doi:10.1128/JVI.00708-14

Title: Investigating a Mystery Disease: Tales from a Viral Detective

Author: W. Ian Lipkin[1]

1. Center for Infection and Immunity, Mailman School of Public Health, Columbia University, New York, NY, USA

Correspondent footnote: Please send journal correspondence including page proofs, to W. Ian Lipkin: wil2001@columbia.edu.

Abstract: Viral outbreak investigation is challenging logistically as well as scientifically. In the context of addressing a fictional emerging viral disease I describe the process of discovery, from the initial report of a problem through discussions of intellectual property and sample management, study design, management, experimental execution and reporting of results.

Abstract Word Count: 50

Manuscript Word Count: 2,409

Outbreak investigation can be as or more challenging logistically than it is scientifically. This piece is written to provide insights into the entire process, from the initial report of a problem through to resolution, in Charlie Chaplin's fictional country of Tomainia (1). To save on production costs I will play myself and the Greek chorus. Casting for Dr. X is still open. Enjoy.

> Date: Wed 18 Nov 2014 9:00
> From promed-ahead@promed.org
> Subject:
> ENCEPHALITIS: TOMAINIA, REQUEST FOR INFORMATION
> ProMED-mail has received a report from a reliable source of rumors of the occurrence of an outbreak of fatal encephalitis in children in Tomainia. Further information from any informed person or organization in the area would be appreciated [C.C].

Situational awareness

Social networks are increasingly important early sources of information on outbreaks. The Program for Monitoring Emerging Diseases (ProMED) founded in 1994, is frequently the first to report an outbreak. Such reports may begin with a request for information or as report with substantial detail curated by a moderator who provides comments that place the report into context (2).

Contact and vetting for feasibility

ProMED reports may be rapidly followed by messages delivered by phone or email from clinicians or basic scientists who have access to data and materials but do not have the resources needed for diagnosis and discovery. In our experience these messages typically come in during the wee hours of the morning. Indeed, I've yet to field an interesting request for assistance during normal business hours. The first step after making contact is to ascertain whether the outbreak is consistent with an infectious disease, the usual suspects have been considered and what type of investigation samples will allow. I've had requests from investigators at reputable institutions to use an oligonucleotide chip to search for evidence of a prion disease. The story was nonetheless intriguing and we followed it to learn that the outbreak represented an autoimmune disorder triggered by exposure to porcine nervous system tissue (3). A common confound is the lack of samples suitable for molecular discovery. Paraffin embedded tissue, for example, may be useful for immunohistochemistry, *in situ* hybridization or even PCR where assays may have better tolerance for degradation; however, we are reluctant to invest in unbiased high-throughput sequencing with materials that have not been stored at -80C or in buffers designed to preserve nucleic acid integrity.

From: Dr. X

Subject: Re: hello

Date: November 19, 2014 at 2:39 AM EDT

To: "W. Ian Lipkin" <wil2001@columbia.edu>

Dear Professor,

There is a challenging problem in northern Tomainia. You may have heard of outbreaks of unexplained encephalitis that annually result in hundreds of deaths of children. I would like to collaborate with you in investigating the causes of these illnesses.

From: "W. Ian Lipkin" <wil2001@columbia.edu>

Subject: Re: hello

Date: November 19, 2014 at 3:02 AM EDT

To: Dr. X

Dear Dr. X.

We will be happy to collaborate with you on this project. Please suggest a good time to reach you by phone or Skype.

Negotiation

Despite the June 2013 United States Supreme Court ruling in the Association for Molecular Pathology vs. Myriad Genetics that naturally occurring DNA cannot be patented (4), many institutions continue to invest in intellectual property associated with the discovery of microbial gene sequences. It has always been our policy to insist that partners in discovery efforts share equally in whatever equity might arise from collaborations. The financial value of these collaborations is unclear; nonetheless, we continue to include language that reinforces this principal as well academic credit in authorships. A wrinkle that may come up at the publication stage is that authors may be

proposed who have had no role in the study design or execution, and may not even have read the manuscript. Although it is premature to discuss the criteria for authorship during the initial conversation it is unwise to leave it until after the manuscript is prepared. Patient confidentiality receives less emphasis in some countries than others. We recommend de-identifying samples to obviate Institutional Review Board concerns.

From: "W. Ian Lipkin" <wil2001@columbia.edu>

Subject: Re: hello

Date: November 19, 2014 at 9:18 AM EDT

To: Dr. X

Dear Dr. X.

Thanks for taking our call. We will need a material transfer agreement signed by officials at our respective institutions as well as you and me. As discussed, intellectual property and authorships will be jointly shared. All samples must be de-identified to protect the confidentiality of patients and their families.

Material Transfer, Sovereignty and the International Health Regulations

The Convention on Biological Diversity was initiated by the United Nations Environment Programme in 1998 to codify that biological diversity is a global asset and that benefits arising from use of such assets (including microbial sequences and strains) should be equitably shared (5). The 2005 International Health Regulations signed by 196 countries across the globe emphasize capacity building as well as transparency and collaborative surveillance. Although a material transfer agreement addressing equity should provide

reassurance against exploitation, we increasingly encounter resistance to sample export. Furthermore, samples from ungulates in some countries cannot be readily transported due to the potential risks to agriculture such as foot and mouth disease virus. Accordingly, we have built a mobile laboratory that allows us to pursue nucleic acid extractions, targeted molecular assays and serology in-country. This laboratory proved invaluable recently in studying the prevalence of MERS-CoV infection dromedaries in Saudi Arabia (6). We also train colleagues in the developing world to promote self sufficiency. Nonetheless, microbial nucleic acid enrichment, library preparation high-throughput sequencing (HTS) and downstream bioinformatics analysis require state-of-the-art infrastructure. In the 5-10 year time frame, therefore, discovery is likely to require a blend of work at the site of the outbreak and in a reference laboratory.

From: Dr. X

Subject: Re: hello

Date: November 20, 2014 at 3:40 AM EDT

To: "W. Ian Lipkin" <wil2001@columbia.edu>

Dear Professor,

Tomanian officials told me that we cannot send human biological materials sample from Tomania abroad. Can we do the work here?

From: "W. Ian Lipkin" <wil2001@columbia.edu>

Subject: Re: hello

Date: November 20, 2014 at 5:33 AM EDT

To: Dr. X

Dear Dr. X,

We will send a team to work with you. However, in the event that PCR fails, I recommend that you join us for a few weeks to finish the workup with high-throughput sequencing.

The hunt begins…

Your colleagues may insist that they have ruled out the usual suspects known to cause disease in a specific catchment area and to ask you to move directly to HTS. Don't do it. More often than not a few simple consensus PCR reactions will solve the mystery. Similarly, although our focus here is on virology, keep an open mind to bacterial, fungal and/or parasitic causes of disease. We have implicated *P. falciparum* in a medical relief worker who died during a Marburg outbreak (7). Additionally, bacteria and viruses may be more pathogenic in concert. During the H1N1 influenza pandemic, the presence of *S. pneumonia* in addition to H1N1 influenza was associated with a more than 100-fold increased risk of severe disease (8). We only appreciated the importance of co-infection after sequencing hundreds of influenza viruses in an effort to understand an increase in morbidity and mortality in Argentina.

Communicate with the clinicians as well as the laboratorians. There are exceptions where application of HTS discovery methods leads to resolution of a single case and a high profile publication. However, these will typically be cases where one implicates an agent that is known to be pathogenic and a simple solution like targeted consensus PCR would have succeeded more rapidly and at lower cost. Clinicians can be helpful in

157 providing a differential diagnosis that may allow such targeted assays. In an effort to
158 identify clusters of disease, laboratorians frequently oversimplify clinical data. It is
159 important to determine the basis by which clusters have been identified. In the ideal
160 circumstance, cases within clusters are clearly related with respect to timing of onset,
161 geography, subject age and syndromic features. However, even in the absence of a
162 classical cluster, one can frequently tease out which samples can be considered
163 together and which should not through detailed discussion with the clinicians who
164 collected the samples. Similarly, to avoid investing in samples that may be
165 contaminated or degraded, it is important to communicate with the laboratorian
166 regarding how samples have been processed and stored. Bear in mind the adage
167 “garbage in, garbage out”. We have been frustrated more than once to learn in
168 retrospect that gloves or dissecting tools were used continuously to process multiple
169 samples from different individuals, resulting in the appearance of a cluster wherein
170 several individuals were infected with the same agent. Another challenge is laboratories
171 in which sample freezers are adjacent to tissue culture facilities where viruses are
172 passaged. One can sort artifact from reality at later steps in the discovery process
173 through collection of additional samples or testing for the presence of viruses within
174 cells using immunohistochemistry or *in situ* hybridization; however, this is time
175 consuming and requires additional resources.
176
177 *Controls*
178 The selection and use of controls in viral discovery projects are critical but complex. In
179 bacterial microbiome projects (I don't think that virologists, mycologists or parasitologists

should cede the term microbiome to bacteriology), we simultaneously profile bacterial populations of cases and controls based on the assumption neither cases nor controls are sterile. We typically pursue viral discovery in cases alone. This is not because controls are sterile, but because the lower number of viral sequences in most clinical samples allows one to more easily sift through them to identify candidates that can be tested for relationship with disease using less expensive and more sensitive methods such as PCR. A more critical step is the selection of controls. There are instances where investigators have collected cases of a disease within a specific catchment area during an outbreak of disease and reported identifying the cause of the disease based on absence of the agent in controls collected during a different season in the same catchment area. It is imperative that controls be selected to ensure that findings in cases are associated with disease rather than variability in virus incidence with season, geography or socioeconomic status.

Fallability of the Willy Sutton rule

The apocryphal Willie Sutton quote (denied by Sutton) that he robbed banks “because that’s where the money is” (Willie Sutton and Edward Linn, Where the Money was: Memoirs of a Bank Robber) can direct sampling when the affected organ is readily accessible as in respiratory, diarrheal, genitourinary and skin diseases or hemorrhagic fevers. However, central nervous system (CNS) diseases and autoimmune disorders may require a less direct approach. Brain biopsies are rare and although we have used postmortem samples to enable efficient viral discovery (9, 10), a post mortem diagnosis is not as gratifying as an ante mortem finding that impacts patient care. Examples of the

latter include the discovery and implication of LuJo virus in a hemorrhagic fever outbreak where use of ribavirin probably saved one victim (11) and of a novel polyomavirus in a transplant recipient wherein assays for viremia facilitated regulation of immunomodulatory therapy, controlling infection while sparing the transplanted organ (12).

Most often the only available CNS sample is cerebrospinal fluid. In bacterial CNS infections, cerebrospinal fluid (CSF) culture and PCR are frequently informative; however, with the exception of herpes encephalitis, the rate for resolution of viral CNS infections is 50% or lower (13). One can increase this success rate by examining oral and fecal samples by PCR for the presence of viruses known to be associated with CNS infection (e.g. enteroviruses) or by testing CSF for antibodies to neurotropic viruses (e.g., West Nile virus). Some cases of what appear to be infectious encephalitis are in fact autoimmune disorders wherein antibodies bind to brain components such as receptors. Although it is important to distinguish the latter as they respond to immunosuppressive rather than antiviral or antimicrobial therapy, these disorders to not present in clusters as they typically represent paraneoplastic syndromes. You will also want serum, preferably collected during the illness and few weeks later, to test for evidence of an immune response to the agent or agents you implicate using molecular methods.

From: "W. Ian Lipkin" <wil2001@columbia.edu>

Subject: Re: hello

Date: January 14, 2014 at 9:47 AM EDT

To: Dr. X

Dear Dr. X,

As discussed, we started with a multiplex PCR assay that targets 20 bacteria and viruses frequently implicated in CNS infections. Five of the 10 children in the July outbreak had enterovirus sequences in their CSF by 5'UTR; 9 had enterovirus sequences in their feces. Although we don't have CSF from normal children, only 2 of 10 had viral sequences in their feces. VP1 sequencing indicates that all children were infected with the same genotype. I think we can be confident that this is the agent. Nonetheless, lets' build a LIPS assay (14) and test for changes in antibody titer in acute and convalescent sera. We can also use this same assay to determine the prevalence of infection in the general population and, if you have samples stored from earlier years, when the infection first emerged in Tomainia. We are less clear in the cases collected in November outbreak. No real clustering… just a mix of what appear to be opportunistic infections in association with HIV--cryptococcus, toxoplasma and Epstein Barr virus. The December cases are particularly interesting. HTS of brain material from the one postmortem specimen after treatment with nucleases to eliminate nucleic acid not protected within nucleocapsids revealed the presence of a novel flavivirus that we have tentatively named Tomainian flavirurus (TFV). *In situ* hybridization confirms the presence of virus in neurons. We created real time PCR and LIPS assays and found that 3 of 5 cases had TFV sequences in their CSF and all 5 had antibodies to TFV in CSF and blood. We've also found serum antibodies in 100 of 500 random controls; thus, it seems that while TFV is not a rare infection, the majority of infections do not result in encephalitis. The similarity

in presentation to West Nile encephalitis led us to investigate mosquito pools using PCR. The vector seems to be *Culex tomainia*. The reservoir is still unknown; however, I've seen reports of dead dodos in Jurassic park. We could wait to close that loop but what we have already is more than sufficient for a high impact publication.

Acknowledgements:

I am grateful for thoughtful comments from Ellie Kahn, Amit Kapoor, Thomas Briese and Mady Hornig.

References:

1. **Chaplin C.** 1940. The Great Dictator. United Artists.
2. **ProMED-mail** 2010. About ProMED-mail. International Society for Infectious Diseases. [Online.]
3. **Holzbauer SM, DeVries AS, Sejvar JJ, Lees CH, Adjemian J, McQuiston JH, Medus C, Lexau CA, Harris JR, Recuenco SE, Belay ED, Howell JF, Buss BF, Hornig M, Gibbins JD, Brueck SE, Smith KE, Danila RN, Lipkin WI, Lachance DH, Dyck PJ, Lynfield R.** 2010. Epidemiologic investigation of immune-mediated polyradiculoneuropathy among abattoir workers exposed to porcine brain. PloS one **5:**e9782.
4. **Supreme Court of the United States.** 2013. Association for Molecular Pathology Et al. v. Myriad Genetics, Inc. Et al., vol. 12-398.

5. **Secretariat of the Convention on Biological Diversity** 2014. Convention on Biological Diversity. [Online.]

6. **Alagaili AN, Briese T, Mishra N, Kapoor V, Sameroff SC, Burbelo PD, de Wit E, Munster VJ, Hensley LE, Zalmout IS, Kapoor A, Epstein JH, Karesh WB, Daszak P, Mohammed OB, Lipkin WI.** 2014. Middle East respiratory syndrome coronavirus infection in dromedary camels in Saudi Arabia. mBio **5:**e00884-00814.

7. **Palacios G, Quan PL, Jabado OJ, Conlan S, Hirschberg DL, Liu Y, Zhai J, Renwick N, Hui J, Hegyi H, Grolla A, Strong JE, Towner JS, Geisbert TW, Jahrling PB, Buchen-Osmond C, Ellerbrok H, Sanchez-Seco MP, Lussier Y, Formenty P, Nichol MS, Feldmann H, Briese T, Lipkin WI.** 2007. Panmicrobial oligonucleotide array for diagnosis of infectious diseases. Emerging infectious diseases **13:**73-81.

8. **Palacios G, Hornig M, Cisterna D, Savji N, Bussetti AV, Kapoor V, Hui J, Tokarz R, Briese T, Baumeister E, Lipkin WI.** 2009. Streptococcus pneumoniae coinfection is correlated with the severity of H1N1 pandemic influenza. PloS one **4:**e8540.

9. **Briese T, Jia XY, Huang C, Grady LJ, Lipkin WI.** 1999. Identification of a Kunjin/West Nile-like flavivirus in brains of patients with New York encephalitis. Lancet **354:**1261-1262.

10. **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.

294 Astrovirus encephalitis in boy with X-linked agammaglobulinemia. Emerging
295 infectious diseases **16:**918-925.
296 11. **Briese T, Paweska JT, McMullan LK, Hutchison SK, Street C, Palacios G,**
297 **Khristova ML, Weyer J, Swanepoel R, Egholm M, Nichol ST, Lipkin WI.** 2009.
298 Genetic detection and characterization of Lujo virus, a new hemorrhagic fever-
299 associated arenavirus from southern Africa. PLoS pathogens **5:**e1000455.
300 12. **Mishra N, Pereira M, Rhodes RH, An P, Pipas JM, Jain K, Kapoor A, Briese**
301 **T, Faust PL, Lipkin WI.** 2014. Identification of a Novel Polyomavirus in a
302 Pancreatic Transplant Recipient With Retinal Blindness and Vasculitic Myopathy.
303 The Journal of infectious diseases.
304 13. **Kennedy PG.** 2005. Viral encephalitis. Journal of neurology **252:**268-272.
305 14. **Burbelo PD, Ching KH, Bush ER, Han BL, Iadarola MJ.** 2010. Antibody-
306 profiling technologies for studying humoral responses to infectious agents. Expert
307 review of vaccines **9:**567-578.

308