

## Opinion

## Pathogen Discovery

W. Ian Lipkin\*

Center for Infection and Immunity, Mailman School of Public Health of Columbia University, New York, New York, United States of America

“What has become clear to you since we last met?”

—Benjamin Franklin

Discovery—literally, the act of uncovering—implies a new and fundamental observation that changes the way in which we view and respond to the world (Image 1). Eureka (“I have found it”) moments are rare in science. Given the physiques of many scientists, this may be a good thing. Legend has it that when Archimedes discovered while bathing that the volume of an object could be calculated by finding the volume of water it displaced, he leaped out of the bathtub and ran naked through the streets of Syracuse proclaiming his discovery.

The rate of discovery of new microbes, and of new associations of microbes with health and disease, has accelerated over the past two decades. Many factors are implicated. New pathogens have truly emerged with the globalization of travel and trade, changes in demographics and land use, susceptibility to opportunistic organisms associated with immunosuppression, and climate change [1]. New



**Image 1. The author engaging in the process of scientific discovery.**

doi:10.1371/journal.ppat.1000002.g001

molecular technologies such as MassTag PCR [2–5], microbial microarrays [6–9], and unbiased high throughput sequencing (HTS) [10] have enabled efficient microbial surveillance and discovery. The databases needed to recognize sequences as host or microbial have improved dramatically. Sample collection has become sophisticated and comprehensive. Last, but not least, our models for pathogenesis embrace increasingly complex mechanisms that consider host–microbe–timing interactions in acute and chronic disease.

### Proof of Causation

Finding an organism is only one step in establishing a causal relationship or understanding how it causes disease. Many have wrestled with the challenge of codifying the process of proving causation. Based on the germ theory of disease of Pasteur, Koch and Loeffler proposed criteria that define a causative relationship between agent and disease: the agent is present in every case of a disease; it is specific for that disease; and it can be propagated in culture and inoculated into a naïve host to cause the same disease [11]. Known as Koch’s postulates, these criteria were modified by Rivers for viruses [12], and by Fredericks and Relman to reflect the introduction of molecular methods [13]. Although fulfillment of Koch’s postulates remains the most persuasive evidence of causation, there are problems with holding to this standard. Overlap in signs and symptoms due to infection with different agents is common. Results of infection may vary with genetic background, age, nutrition, and previous exposure to similar agents. Many agents cannot be cultured; furthermore, there may be no animal model. Proving causa-

tion is particularly challenging where agents have remote effects or require cofactors for expression. In many acute infectious diseases, the responsible agent is readily implicated because it replicates at high levels in the affected tissue at the time the disease is manifest, morphological changes consistent with infection are evident, the agent is readily identified with classical or molecular methods, and there is evidence of an adaptive immune response. However, implication is more difficult when classical hallmarks of infection are absent or mechanisms of pathogenesis are indirect or subtle. Here, one may resort to a statistical assessment of the strength of epidemiological association based on the presence of the agent or its footprints (nucleic acid, antigen, and preferably, an immune response), and biological plausibility as indicated by analogy to diseases with other organisms where linkage is persuasive.

### Pathways to Pathogenesis

Implication of an agent is easiest if it is present in high concentration at the site of disease when the disease is manifest. Examples include poliomyelitis, where death of infected motor neurons results in paralysis, or infectious diarrheas where the causative agent (bacterium, virus, or parasite) is found in the gastrointestinal tract. More complex examples of intoxication occur in botulism or tetanus, where replication in the subcutaneous tissues or the gastrointestinal tract results in release of toxins that have remote effects on the nervous system. Pathogenesis may be immune-mediated as in hepatocellular carcinoma due to persistent hepatitis B or hepatitis C infections [14,15]. Infection can also impair immune function, opening

**Citation:** Lipkin WI (2008) Pathogen Discovery. *PLoS Pathog* 4(4): e1000002. doi:10.1371/journal.ppat.1000002

**Editor:** Marianne Manchester, The Scripps Research Institute, United States of America

**Published:** April 25, 2008

**Copyright:** © 2008 W. Ian Lipkin. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** Support for our work comes from NIH awards AI070411, HL83850, and NS047537, and U54AI5758 (Northeast Biodefense Center-Lipkin).

**Competing Interests:** The author has declared that no competing interests exist.

\* E-mail: wil2001@columbia.edu

the door to opportunistic pathogens. This observation dates back to the early 1900s when von Pirquet noted the loss of skin reactivity to tuberculin in association with measles infection [16]; however, it is now best known in the context of HIV/AIDS. The effects of infection may depend on the age and maturation status of the host. Individuals at either extreme of life are at increased risk for morbidity and mortality [17–19]. Infection during organogenesis may have different consequences than at other times. Birth defects can accompany prenatal infection with toxoplasma, rubella-, cytomegalo-, and herpesviruses [20]. Persistent viral infections can have subtle effects on cellular physiology that result in alterations in the expression of neurotransmitters or hormones that have profound effects, including cognitive impairment [21], hypothyroidism [22], or diabetes mellitus [23]. Infection can break tolerance for “self,” resulting in autoimmune disease [24]. Autoimmunity may be restricted to the tissue in which the agent replicates, presumably because host antigens in that tissue are presented in a new context. However, cross-reactive immune responses may have an impact at a distance as in molecular mimicry, where streptococcal infection in the pharynx or the skin results in immunity to the heart and brain, causing valvular disease and chorea, respectively [25,26]. The constellation of direct and indirect pathways to disease, short and long term, poses challenges for pathogen discovery.

## Strategies for Pathogen Discovery

Although reviews on pathogen surveillance and discovery typically focus on the latest molecular technologies, it is important to emphasize the pivotal roles of clinical acumen, pathology, serology, and classical culture techniques. Clinicians and epidemiologists are the prime movers in pathogen discovery. They recognize the appearance of new syndromes, collect materials for investigation, and persuade their colleagues to take up the search. Anatomic pathology can be instrumental in directing molecular work. The discoveries of Nipah virus [27] and West Nile virus [28–31] were facilitated by demonstration of henipavirus and flavivirus proteins in tissues, which allowed focused consensus PCR analyses. Classical virological methods such as tissue culture and

serology proved pivotal in the 2003 SARS outbreak [32]. Propagation of a virus in tissue culture enabled its rapid characterization by consensus and random PCR cloning, microarray, and electron microscopy.

The advent of methods for cloning nucleic acids of microbial pathogens directly from clinical specimens ushered in a new era in pathogen discovery. Over the past two decades, subtractive cloning, expression cloning, consensus PCR, and high throughput pyrosequencing resulted in identification of novel agents associated with both acute and chronic diseases, including Borna disease virus, hepatitis C virus, Sin Nombre virus, HHV-6, HHV-8, *Bartonella henselae*, *Tropheryma whippelii*, Nipah virus, SARS coronavirus, and Israel Acute Paralysis virus [10,27,32–40].

The most familiar molecular assays in clinical microbiology are singleplex PCR assays designed to detect and quantitate the burden of individual candidate organisms. These have revolutionized blood banking, drug management of HIV, and containment of outbreaks of infectious disease. Degenerate primers can be employed in singleplex PCR assays at reduced stringency to facilitate detection of related but unknown organisms. However, clinical manifestations are not typically pathognomonic of infection with specific pathogens; thus, unless an investigator has clues to narrow a search, this is a cumbersome strategy even if samples, resources, and time are sufficient to invest in many singleplex assays for different agents.

In contrast, multiplex assays simultaneously entertain many hypotheses. The number of candidates considered can range from less than ten with multiplex PCR, to thousands with microarrays, to the entire tree of life with HTS. Costs and ease of use are improving; nonetheless, only multiplex PCR assays are widely used at present. In microarrays and HTS, many genetic targets compete for assay components (e.g., nucleotides, polymerases, and dyes), in some instances with different efficiencies. This abrogates quantitation and reduces sensitivity.

## A Staged Strategy for Pathogen Detection and Discovery

Costs of pathogen discovery can be contained by staging investment. Where epidemiology, serology, or pathology sug-

gest one or a few candidates, singleplex PCR is ideal. Where no such clues pertain or singleplex assays fail, syndromic multiplex PCR assays allow rapid examination of up to 30 candidates with only a modest increase in time and expense. Microarrays are indicated if multiplex PCR fails. Each of these methods requires that an agent be related to one already known. Agents novel or sufficiently distant in sequence to confound hybridization may require subtractive cloning or HTS. Irrespective of the route that results in identification of a candidate, subsequent steps include quantitation of pathogen burden in affected hosts and controls, detailed characterization of the pathogen for features that may contribute to virulence or provide clues to provenance, and serology as an index to acute infection and as a tool to examine prevalence of infection over time and geography.

## Future Perspectives

Nucleic acid platforms are continually evolving, enabling faster, more sensitive, and less expensive methods for direct microbial detection. Improvements in development for these platforms include microfluidic sample processing and direct measurement of conductance changes with hybridization. Proteomics and host response profiling may yet yield diagnostic biomarkers. The most advanced technology will fail if samples are degraded, and data will be uninterpretable without accurate information on clinical course and sample provenance; thus, emphasis should be placed on engaging clinicians as equal partners. In chronic diseases, early exposure and/or genetic susceptibility may influence pathogenesis. The most substantive advances in linking microbes to disease may come from investments in prospective serial sample collections and models wherein diseases reflect intersections of genes and environment in a temporal context.

## Acknowledgments

The author is fortunate to have enjoyed the insights and efforts of many colleagues at The Scripps Research Institute, University of California, Columbia University, and WHO through 20 years of thick (but mostly thin) adventures in pathogen discovery. Thomas Briese, Mady Hornig, Hilary Koprowski, and Josh Lederberg helped develop the concepts presented here.

## References

1. Morse SS (1995) Factors in the emergence of infectious diseases. *Emerg Infect Dis* 1: 7–15.
2. Briese T, Palacios G, Kokoris M, Jabado O, Liu Z, et al. (2005) Diagnostic system for rapid and sensitive differential detection of pathogens. *Emerg Infect Dis* 11: 310–313.

3. Palacios G, Briese T, Kapoor V, Jabado O, Liu Z, et al. (2006) MassTag polymerase chain reaction for differential diagnosis of viral hemorrhagic fever. *Emerg Infect Dis* 12: 692–695.
4. Lamson D, Renwick N, Kapoor V, Liu Z, Palacios G, et al. (2006) MassTag polymerase-chain-reaction detection of respiratory pathogens, including a new rhinovirus genotype, that caused influenza-like illness in New York State during 2004–2005. *J Infect Dis* 194: 1398–1402.
5. Renwick N, Schweiger B, Kapoor V, Liu Z, Villari J, et al. (2007) A recently identified rhinovirus genotype is associated with severe respiratory-tract infection in children in Germany. *J Infect Dis* 196: 1754–1760.
6. Wang D, Coscoy L, Zylberberg M, Avila PC, Boushey HA, et al. (2002) Microarray-based detection and genotyping of viral pathogens. *Proc Natl Acad Sci U S A* 99: 15687–15692.
7. Chiu CY, Rouskin S, Koshy A, Urisman A, Fischer K, et al. (2006) Microarray detection of human parainfluenzavirus 4 infection associated with respiratory failure in an immunocompetent adult. *Clin Infect Dis* 43: e71–e76.
8. Palacios G, Quan P-L, Jabado O, Conlan S, Hirschberg D, et al. (2007) Panmicrobial oligonucleotide array for diagnosis of infectious diseases. *Emerg Infect Dis* 13: 73–81.
9. Quan PL, Palacios G, Jabado OJ, Conlan S, Hirschberg DL, et al. (2007) Detection of respiratory viruses and subtype identification of influenza A viruses by GreeneChipResp oligonucleotide microarray. *J Clin Microbiol* 45: 2359–2364.
10. Cox-Foster DL, Conlan S, Holmes EC, Palacios G, Evans JD, et al. (2007) A metagenomic survey of microbes in honey bee colony collapse disorder. *Science* 318: 283–287.
11. Koch R (1891) Ueber bakteriologische Forschung, Verhandl. des X. Interna. Med. Congr., Berlin 1890, August Hirschwald, Berlin. 35 p.
12. Rivers TM (1937) Viruses and Koch's postulates. *J Bacteriol* 33: 1–12.
13. Fredricks DN, Relman DA (1996) Sequence-based identification of microbial pathogens: a reconsideration of Koch's postulates. *Clin Microbiol Rev* 9: 18–33.
14. Spengler U, Nattermann J (2007) Immunopathogenesis in hepatitis C virus cirrhosis. *Clin Sci (Lond)* 112: 141–155.
15. Levy GA, Chisari FV (1981) A proposed role for the immune system in the pathogenesis of hepatitis B virus induced liver disease. *Crit Rev Clin Lab Sci* 15: 335–353.
16. von Pirquet C (1908) Verhalten der kutanen Tuberkulin-Reaktion während der Masern. *Dtsch Med Wochenschr* 34: 1297–1300.
17. Haeny M (1994) Infection determinants at extremes of age. *J Antimicrob Chemother* 34 Suppl A: 1–9.
18. Gerba CP, Rose JB, Haas CN (1996) Sensitive populations: who is at the greatest risk? *Int J Food Microbiol* 30: 113–123.
19. Ginaldi L, Loreto MF, Corsi MP, Modesti M, De Martinis M (2001) Immunosenescence and infectious diseases. *Microbes Infect* 3: 851–857.
20. Bale JF Jr, Murph JR (1992) Congenital infections and the nervous system. *Pediatr Clin North Am* 39: 669–690.
21. Lipkin WI, Battenberg EL, Bloom FE, Oldstone MB (1988) Viral infection of neurons can depress neurotransmitter mRNA levels without histologic injury. *Brain Res* 451: 333–339.
22. Klavinskis LS, Oldstone MB (1987) Lymphocytic choriomeningitis virus can persistently infect thyroid epithelial cells and perturb thyroid hormone production. *J Gen Virol* 68 (Pt 7): 1867–1873.
23. Oldstone MB, Nerenberg M, Southern P, Price J, Lewicki H (1991) Virus infection triggers insulin-dependent diabetes mellitus in a transgenic model: role of anti-self (virus) immune response. *Cell* 65: 319–331.
24. Fujinami RS, von Herrath MG, Christen U, Whitton JL (2006) Molecular mimicry, bystander activation, or viral persistence: infections and autoimmune disease. *Clin Microbiol Rev* 19: 80–94.
25. Gibofsky A, Kerwar S, Zabriskie JB (1998) Rheumatic fever. The relationships between host, microbe, and genetics. *Rheum Dis Clin North Am* 24: 237–259.
26. Husby G, van de Rijn I, Zabriskie JB, Abidin ZH, Williams RC Jr (1976) Antibodies reacting with cytoplasm of subthalamic and caudate nuclei neurons in chorea and acute rheumatic fever. *J Exp Med* 144: 1094–1110.
27. Paton NI, Leo YS, Zaki SR, Auchus AP, Lee KE, et al. (1999) Outbreak of Nipah-virus infection among abattoir workers in Singapore. *Lancet* 354: 1253–1256.
28. 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.
29. Lanciotti RS, Roehrig JT, Deubel V, Smith J, Parker M, et al. (1999) Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States. *Science* 286: 2333–2337.
30. Shieh WJ, Guarner J, Layton M, Fine A, Miller J, et al. (2000) The role of pathology in an investigation of an outbreak of West Nile encephalitis in New York, 1999. *Emerg Infect Dis* 6: 370–372.
31. Steele KE, Linn MJ, Schoepp RJ, Komar N, Geisbert TW, et al. (2000) Pathology of fatal West Nile virus infections in native and exotic birds during the 1999 outbreak in New York City, New York. *Vet Pathol* 37: 208–224.
32. Peiris JS, Lai ST, Poon LL, Guan Y, Yam LY, et al. (2003) Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* 361: 1319–1325.
33. Challoner PB, Smith KT, Parker JD, MacLeod DL, Coulter SN, et al. (1995) Plaque-associated expression of human herpesvirus 6 in multiple sclerosis. *Proc Natl Acad Sci U S A* 92: 7440–7444.
34. Chang Y, Cesarman E, Pessin MS, Lee F, Culpepper J, et al. (1994) Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma. *Science* 266: 1865–1869.
35. Choo QL, Kuo G, Weiner AJ, Overby LR, Bradley DW, et al. (1989) Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 244: 359–362.
36. Lipkin WI, Travis GH, Carbone KM, Wilson MC (1990) Isolation and characterization of Borna disease agent cDNA clones. *Proc Natl Acad Sci U S A* 87: 4184–4188.
37. Nichol ST, Spiropoulou CF, Morzunov S, Rollin PE, Ksiazek TG, et al. (1993) Genetic identification of a hantavirus associated with an outbreak of acute respiratory illness. *Science* 262: 914–917.
38. Relman DA, Loutit JS, Schmidt TM, Falkow S, Tompkins LS (1990) The agent of bacillary angiomatosis. An approach to the identification of uncultured pathogens. *N Engl J Med* 323: 1573–1580.
39. Relman DA, Schmidt TM, MacDermott RP, Falkow S (1992) Identification of the uncultured bacillus of Whipple's disease. *N Engl J Med* 327: 293–301.
40. VandeWoude S, Richt JA, Zink MC, Rott R, Narayan O, et al. (1990) A borna virus cDNA encoding a protein recognized by antibodies in humans with behavioral diseases. *Science* 250: 1278–1281.