Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Just when we think we have a good understanding of the microbes that occupy our planet and cause disease, nature reminds us that our knowledge is incomplete. In the last few decades, a novel retrovirus (HIV) has emerged from the jungles of Africa to infect and kill millions of individuals from all continents. Less dramatically, newly described microbes have been linked to human diseases since 1980 (Figure 1). Are there other microbial pathogens that are significant causes of human disease and await discovery? There are many diseases for which a microbial aetiology is suspected (Figure 2). The hypothesis that a disease has an infectious cause is supported by:

- clinical features – similar to those of known infectious diseases (e.g. fever, leucocytosis)
- epidemiology – case clustering in time or location
- histology – inflammation of affected tissues (e.g. granulomata) or characteristic microbial structures
- treatment – a clinical response to antimicrobial treatment
- prevention of disease by vaccines targeting microbial antigens.

Proof that a microbe causes a disease requires more rigorous evidence. If microbes are the cause of some idiopathic diseases, why have we not discovered them yet?

- One reason is failure to consider the hypothesis of infection. The association between Helicobacter pylori infection and peptic ulcer disease was not made until recently, despite our ability to cultivate this bacterium in the laboratory and to see it under the microscope in gastric biopsies.
- Another reason is failure of conventional microbial detection technologies. Studies of bacterial biodiversity in various environmental and human niches have shown that cultivation can detect only a fraction of the bacteria identified using nucleic acid sequences. Microbes may exist in viable but non-culturable states, or as sessile members of biofilm communities in which individual cells may be difficult to cultivate. Accordingly, failure to propagate a microbe in the laboratory does not mean that the sample is free of microbes.
- Microbes may evade detection through a pathogenic process of ‘hit and run’. Organisms such as Streptococcus pyogenes may initiate an immunological response that continues to cause disease long after the microbe is eradicated (rheumatic fever).
- Ubiquitous microbes can cause rare diseases as a result of aberrant host responses. When a common virus causes a rare neurological disease in a small subset of susceptible hosts, for example, it is difficult to make an association. Infection with the

**Some micro-organisms described since 1980 and the diseases they cause**

- HIV
- Sin Nombre virus
- Hepatitis C virus
- Human herpesvirus 8
- Nipah virus
- SARS coronavirus
- *Escherichia coli* O157:H7
- *Helicobacter pylori*
- *Bartonella henselae*
- *Tropheryma whippelii*
- *Ehrlichia chaffeensis*
- *Anaplasma phagocytophilum*
- *Borrelia burgdorferi*
- *Cyclospora cayatenensis*

1

**Idiopathic diseases with a suspected infectious cause**

- Crohn’s disease
- Ulcerative colitis
- Bacterial diarrhoea
- Multiple sclerosis
- Kawasaki disease
- Rheumatoid arthritis
- Wegener’s granulomatosis
- Kikuchi–Fugimoto disease
- Sarcoidosis
- Atherosclerosis
- Diabetes mellitus
- Schizophrenia

2
virus is not a helpful risk factor in predicting development of disease because exposure is universal.

- Classic epidemiological clues suggesting microbial participation may be missing for some chronic infectious diseases (e.g. Whipple’s disease).
- We tend to believe the paradigm that one microbe causes one disease. However, disease may result from changes in an entire microbial community and may not be caused by a single microbe; periodontitis and bacterial vaginosis are examples. Efforts to study such diseases must assess the dynamic interaction between microbes.

**Sequence-based approaches**

Many of the recently discovered microbial pathogens were identified by detecting their unique nucleic acid sequences in tissues. An example of this approach is the discovery of *Bartonella* as the cause of bacillary angiomatosis. Although bacterial structures are visible in bacillary angiomatosis tissues, a cultivated bacterium was not initially apparent. Oligonucleotide primers complementary to highly conserved sequences in known bacterial 16S rRNA genes were used in a polymerase chain reaction (PCR) to amplify small quantities of bacterial DNA present in bacillary angiomatosis tissues. The 16S rRNA gene also contains regions of sequence variability that, when amplified using this consensus PCR approach, can be used to identify the bacteria or to infer evolutionary relationships to other bacteria.

A unique 16S rDNA sequence was detected in the tissues of three patients with bacillary angiomatosis but not in control tissues, identifying a bacterium in the *Bartonella* genus as a cause of the disease. This organism was later named *Bartonella henselae*. *B. quintana* is another cause of bacillary angiomatosis, and also causes trench fever and endocarditis, and *B. henselae* is responsible for cat-scratch disease in immunocompetent hosts. *B. bacilliformis* (Oroya fever, verruga peruana) and *B. elizabethae* (endocarditis) are also human pathogens.

A sequence-based approach to novel pathogen discovery has several advantages. All infectious agents should be detectable because they contain nucleic acids (excluding prions), and because each contains a unique complement of nucleic acid, different microbes should be distinguishable. Nucleic acid amplification and probe-based signal amplification methods allow detection of microbial nucleic acid sequences down to the single-copy level, yielding impressive assay sensitivities.

Several sequence-based approaches to pathogen discovery have yielded novel microbes that cause human disease. Consensus PCR has been used to identify the bacterial cause of Whipple’s disease (†*Tropheryma whippelii*) and the viral cause of hantavirus pulmonary syndrome (†*Sin Nombre virus* and others). Hepatitis C virus was identified by screening expression libraries of cDNA for antigen that reacted with immune serum from patients. Human herpesvirus 8 (the cause of Kaposi’s sarcoma) was identified using representational difference analysis. In this method, a unique microbial DNA fragment was enriched in a tissue sample using subtractive hybridization and PCR-based amplification.

The novel coronavirus associated with severe acute respiratory syndrome (SARS) was detected using a pan-viral DNA microarray. An unidentified virus was cultivated from a patient with SARS using Vero cells, the RNA from this culture was reverse transcribed and amplified using PCR, and the labelled cDNA was hybridized to a microarray containing oligomers complementary to numerous known viruses. Hybridization of the cDNA to known coronavirus sequences suggested that the viral isolate was a member of the coronavirus family. The viral genome revealed that it was a newly characterized coronavirus.

**The future**

Future attempts to identify novel microbes associated with human disease may use other sequence-based approaches. High-throughput sequencing may allow identification of unique microbial nucleic acid sequences in a background of host DNA. The complete sequencing of the human genome and multiple microbial genomes make this approach more feasible.

DNA microarrays are also likely to be used in the search for novel pathogens. Microarrays containing thousands of DNA spots, representing a diversity of microbial and viral genes, may be used to detect microbial nucleic acids in tissues by hybridization. Microarrays of human DNA may also be used to monitor host gene expression in disease, as a means of looking for characteristic host response profiles that help identify infectious agents. Identification of novel or previously described infectious agents may improve diagnosis, prevention and treatment of disease. Because of the continued evolution of microbes and of humans, the emergence of new pathogens and diseases is a virtual certainty.

**REFERENCES**

1. Relman D A, Loutit J S, Schmidt T M et al. The agent of bacillary angiomatosis: an approach to the identification of uncultured pathogens. *N Engl J Med* 1990; 323: 1573–80.
2. Relman D A, Schmidt T M, MacDermott R P et al. Identification of the uncultured bacillus of Whipple’s disease. *N Engl J Med* 1992; 327: 293–301.
3. Nichol S T, Spiropoulou C F, Morzunov S et al. Genetic identification of a hantavirus associated with an outbreak of acute respiratory illness. *Science* 1993; 262: 914–17.
4. Choo Q L, Kuo G, Weiner A J et al. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 1989; 244: 359–62.
5. Chang Y, Cesarmian E, Pessin M S et al. Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi’s sarcoma. *Science* 1994; 266: 1865–9.
6. Wang D, Urisman A, Liu Y T et al. Viral discovery and sequence recovery using DNA microarrays. *PLoS Biol* 2005; 1: E2.

**FURTHER READING**

Boldrick J C, Alizadeh A A, Diehn M et al. Stereotyped and specific gene expression programs in human innate immune responses to bacteria. *Proc Natl Acad Sci U S A* 2002; 99: 972–7.

Fredricks D N, Relman D A. Infectious agents and the etiology of chronic idiopathic diseases. *Curr Clin Top Infect Dis* 1998; 18: 180–200.

Kroes I, Lepp P W, Relman D A. Bacterial diversity within the human subgingival crevice. *Proc Natl Acad Sci U S A* 1999; 96: 14,547–52.

Pace N R. A molecular view of microbial diversity and the biosphere. *Science* 1997; 276: 734–40.

Rota P A, Oberste M S, Monroe S S et al. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. *Science* 2003; 300: 1394–9.