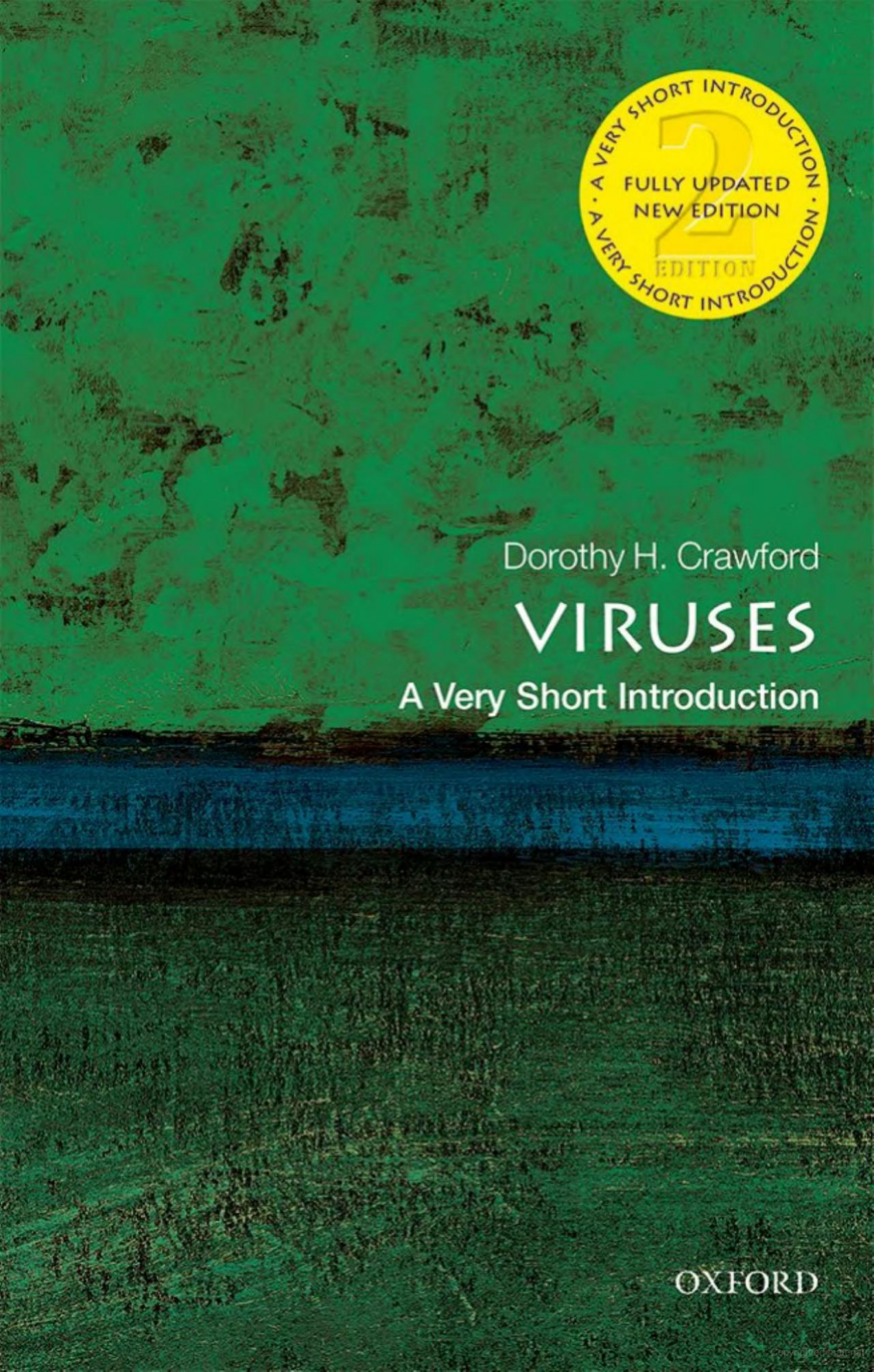

A VERY SHORT INTRODUCTION
FULLY UPDATED
NEW EDITION
2
EDITION
A VERY SHORT INTRODUCTION
Dorothy H. Crawford
VIRUSES
A Very Short Introduction
OXFORD

Dorothy H. Crawford

VIRUSES

A Very Short Introduction

SECOND EDITION

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

First edition published 2011
Second edition published 2018

Impression: 1

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data

Data available

Library of Congress Control Number: 2017961312

ISBN 978–0–19–881171–8

Printed in Great Britain by
Ashford Colour Press Ltd, Gosport, Hampshire

Contents

Acknowledgements

My thanks go to Dr Ingolfur Johannessen for his professional advice.

I am deeply indebted to Dr Karen McAulay who provided essential information for the section of the book on arthropod-transmitted viruses that is new to this edition.

List of illustrations

Roser. <https://ourworldindata.org/wp-content/uploads/2013/11/HIV-AIDS.png>

11 The yellow fever transmission cycle **50**

From D. H. Crawford, *The Invisible Enemy* (OUP, 2000), p. 26, fig. 1.4 © Oxford University Press.

12 The worldwide distribution of dengue fever, Zika, and Chikungunya viruses **52–4**

Jessica Patterson, MD et al., 'Dengue, Zika and Chikungunya: Emerging Arboviruses in the New World', *Western Journal of Emergency Medicine*, Vol. 17, Issue 6 (November 2016).

13 Croup hospitalizations in children under 15 in the USA, 1981–2002 **69**

Reproduced from S. D. Roxborgh et al, 'Trends in pneumonia and empysema in Scottish children in the past 25 years', BMJ Vol. 93 (April 1, 2008.) with permission from BMJ publishing Group Ltd.

14 CD4 count and viral load during acute, asymptomatic, and symptomatic phases of HIV infection **85**

Reproduced from A. Mindel and M. Tenant-Flowers, 'Natural History and Management of early HIV infection', ABC of Aids (2001), with permission from BMJ Publishing Group Ltd.

15 World map showing the prevalence of HBV and HCV infections **89**

Based on WHO data.

16 World map showing the prevalence of HTLV-1 infection **97**

Based on WHO data.

17 Burkitt's map of the distribution of Burkitt's lymphoma in Africa **101**

Reproduced from D. Burkitt, 'Determining the Climatic Limitations of Children's Cancer Common in Africa', British Medical Journal, 2 (1962): 1019–23, with permission from BMJ Publishing Group.

18 Age-standardized incidence and mortality rates for cervical cancer by world region, 2002 estimates **108**

Reproduced with permission from Globocan 2002: Cancer Incidence Mortality and Prevalence Worldwide IARC CancerBase No. 5, version 2.0, 2004, Lyon: IARCPress. With permission from IARC, 2017.

19 *The Cow-Pock-or-The Wonderful Effects of the New Inoculation*, by James Gillray, 1802 **112**

British Cartoon Prints Collection. Prints & Photographs Division, Library of Congress, LC-USZC4-3147.

Introduction

In December 2019 a novel virus jumps to humans from an unspecified animal in a live animal market in Wuhan, China. Despite rapid lock-down of the Wuhan area and a travel ban, within weeks the virus spreads throughout China, infecting thousands. Identified as a coronavirus related to SARS, the airborne virus, named COVID-19, causes a respiratory disease, with a fatality rate of around 2%. By 30th January it reaches over 20 other countries by hitching a ride with infected air-travellers. Health services worldwide quarantine suspected cases and WHO declares a public health emergency of international concern. Nevertheless, the virus establishes chains of infection in countries in both Asia and Europe and the death toll rises. By March 2020, as health services struggle to contain the outbreaks, worldwide spread seems inevitable.

In this book the author introduces viruses and their lifestyle, the constant battle between viruses and the immune system of infected individuals, and how new viruses like the 2019 coronavirus emerge and cause outbreaks, epidemics, and pandemics. Later chapters consider viruses that persist in the body for life, some of which can cause tumours. The book concludes with a historical look at the changing pattern of virus infections, and speculates on future challenges they present. A glossary explains specialist terms and the derivation of virus names.

Chapter 1
What are viruses?

The microbe is so very small
You cannot make him out at all,
But many sanguine people hope
To see him through a microscope.
His jointed tongue that lies beneath
A hundred curious rows of teeth;
His seven tufted tails with lots
Of lovely pink and purple spots,
On each of which a pattern stands,
Composed of forty separate bands;
His eyebrows of a tender green;
All these have never yet been seen—
But Scientists, who ought to know,
Assure us that they must be so…
Oh! let us never, never doubt
What nobody is sure about.

'The Microbe' (1896),
Hilaire Belloc

Primitive microbes evolved on Earth approximately three billion years ago but were isolated by humans only in the late 19th century, around twenty years before Hilaire Belloc wrote 'The Microbe'. Written to amuse, the poem nonetheless reflects the scepticism of

the times. It must have taken a huge leap of faith for people to accept that tiny, living organisms were responsible for diseases that had hitherto been attributed variously to the will of the gods, the alignment of planets, or miasmic vapours emanating from swamps and decomposing organic material. Of course, this realization did not dawn overnight, but as more and more microbes were identified, the 'germ theory' took hold, and by the beginning of the 20th century it was widely accepted even in non-scientific circles that microbes could cause disease.

Key to this momentous leap in understanding were technical developments in microscopes made by the Dutch lens-maker Antonie van Leeuwenhoek (1632–1723) in the 16th century. He was the first to visualize microbes, but it was not until the mid-1800s that Louis Pasteur (1822–95) working in Paris and Robert Koch (1843–1910) in Berlin established 'germs' as the cause of infectious diseases. Pasteur was instrumental in dispelling the general belief in 'spontaneous generation', that is, the generation of life from inorganic material. He showed that mould growth could be prevented in broth by boiling and then placing it in a chamber with filters to exclude the entry of any particulate material from the air. This demonstrated the existence of airborne microscopic 'germs'.

Koch discovered the first bacterium, *Bacillus anthraci*, in 1876. He soon developed methods for growing microbes in the laboratory and then, one after another, the causative microbes of feared diseases like anthrax, tuberculosis, cholera, diphtheria, tetanus, and syphilis were identified and characterized. It became clear that bacteria have a structure similar to mammalian cells with a cell wall surrounding cytoplasm that contains a single, coiled, circular molecule of DNA. The majority are free living and can manufacture all the proteins they need to metabolize, and divide.

However, there remained a group of infectious diseases for which causative organisms could not be found, such as smallpox,

measles, mumps, rubella, and flu. These microbes were obviously very small as they passed through filters that trapped bacteria, and in consequence were called 'filterable agents'. At the time, most scientists thought these were just tiny bacteria.

In 1876, Adolf Mayer (1843–1942), director of the Agricultural Experimental Station in Wageningen, Holland, began to investigate a new disease of tobacco plants which devastated the Dutch tobacco industry. He called it 'tobacco mosaic disease' because of the mottled pattern it produces on the diseased plant's leaves. He showed that the disease was infectious by transmitting it to healthy plants using sap extracted from a diseased one. He thought that the disease was caused by a very small bacterium or a toxin.

Later, biologist Dmitry Ivanovsky (1864–1920) also worked on tobacco mosaic disease at the University of St Petersburg in Russia, and in 1892 demonstrated that its causative agent passed through filters that trapped bacteria. Like Mayer, he suggested that it was caused by a chemical toxin produced by a bacterium.

In 1898, Martinus Beijerinck (1851–1931), from the Agricultural School in Wageningen, followed up on Mayer's experiments by showing that the agent grew in dividing cells and regained its full strength each time it infected a plant. He concluded a living microbe was responsible, and was the first to coin the name *virus*, from the Latin meaning a poison, venom, or slimy fluid.

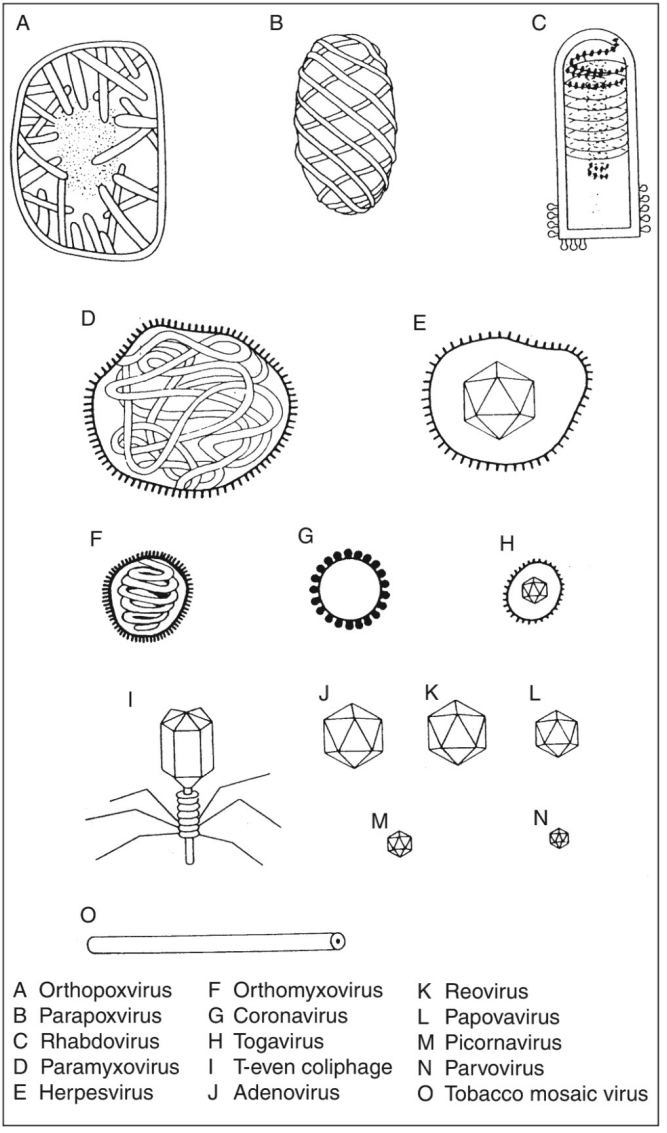
By the beginning of the 20th century, viruses were defined as a group of microbes that were infectious, filterable, and required living cells for their propagation, but the nature of their structure remained a mystery. In the 1930s, tobacco mosaic virus was obtained in crystalline form, suggesting that viruses were purely composed of protein, but shortly afterwards a nucleic acid component was discovered and shown to be essential for infectivity. However, it was not until the invention of the electron microscope

in 1939 that viruses were first visualized and their structure elucidated, showing them to be a unique class of microbes.

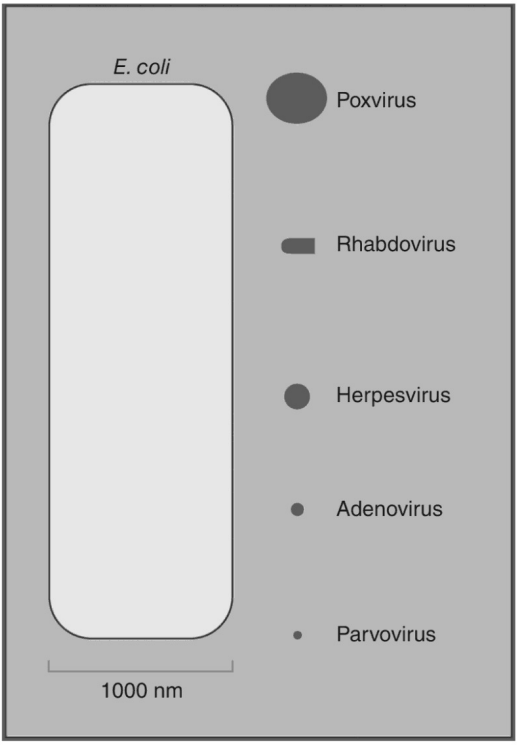
Viruses are not cells but particles. They consist of a protein coat which surrounds and protects their genetic material, or, as the famous immunologist Sir Peter Medawar (1915–87) termed it, 'a piece of bad news wrapped up in protein'. The whole structure is called a *virion* and the outer coat is called the *capsid*. Capsids come in various shapes and sizes, each characteristic of the virus family to which it belongs. They are built up of protein subunits called *capsomeres* and it is the arrangement of these around the central genetic material that determines the shape of the virion. For example, pox viruses are brick-shaped, herpes viruses are icosahedral (twenty-sided spheres), the rabies virus is bullet-shaped, and the tobacco mosaic virus is long and thin like a rod (Figure 1). Some viruses have an outer layer surrounding the capsid called an *envelope*.

Most viruses are too small to be seen under a light microscope. In general, they are around 100 to 500 times smaller than bacteria, varying in size from 20 to 300 nanometres in diameter (nm; 1 nm is a thousand millionth of a metre) (Figure 2). However, the recently discovered giant, the mimivirus (short for 'microbe-mimicking virus'), is an exception, with a diameter of around 700 nm; larger than some bacteria. Inside the virus capsid is its genetic material, or *genome*, which is either RNA or DNA depending on the type of virus (Figure 3). The genome contains the virus's genes, which carry the code for making new viruses, and transmits these inherited characteristics to the next generation. Viruses usually have between 4 and 200 genes, but again mimivirus is most unusual in having an estimated 600 to 1,000 genes, even more than many bacteria.

Cells of free-living organisms, including bacteria, contain a variety of organelles essential for life such as ribosomes that manufacture proteins, mitochondria, or other structures that generate energy,

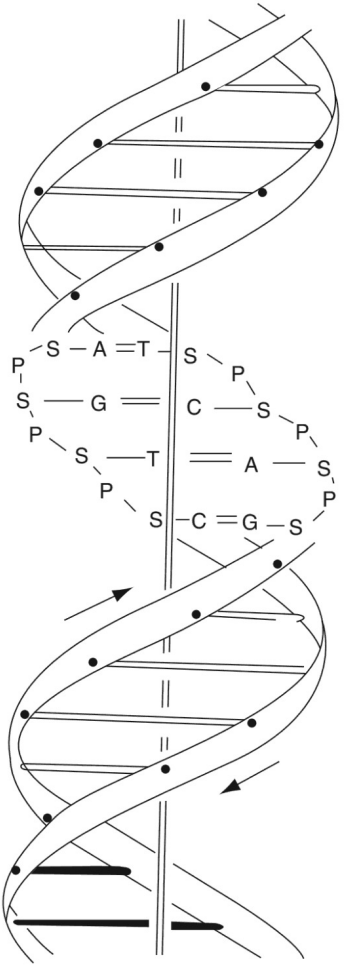


1. The structure of viruses.



2. The comparative sizes of a typical bacterium and representative viruses.

and complex membranes for transporting molecules within the cell, and also across the cell wall. Viruses, not being cells, have none of these and are therefore inert until they infect a living cell. Then they hijack a cell's organelles and use what they need, often killing the cell in the process. Thus viruses are obliged to obtain essential components from other living things to complete their life cycle and are therefore called *obligate parasites*. Even mimivirus, which infects amoebae, has to borrow the amoeba's organelles to manufacture its proteins in order to assemble new mimiviruses.



3. The structure of DNA, showing the two complementary strands that form the helix. The backbone of each strand is composed of molecules of the sugar deoxyribose (S) that are linked to each other through phosphate molecules (P). Each sugar is connected to a nucleotide molecule, and these form the 'letters' of the genetic alphabet. These are: adenine (A), guanine (G), cytosine (C), and thymine (T). The structure of RNA is similar to DNA but its nucleotides are adenine, guanine, cytosine, and uracil (U).

Plant viruses either enter cells through a break in the cell wall or are injected by a sap-sucking insect vector like aphids. They then spread very efficiently from cell to cell via plasmodesmata, pores that transport molecules between cells. In contrast, animal viruses infect cells by binding to specific cell surface receptor molecules. The cell receptor is like a lock, and only viruses that carry the right receptor-binding key can open it and enter that particular cell. Receptor molecules differ from one type of virus to another; although some are found on most cells, others are restricted to certain cell types. A well-known example is the AIDS virus, the human immunodeficiency virus (HIV) that carries the entry key for the CD4 lock, so only cells with CD4 molecules on their surface can be infected by HIV. This specific interaction defines the outcome of the infection, and in the case of HIV leads to destruction of CD4-positive 'helper' T cells that are critical to the immune response. Eventually the immune system fails, and opportunistic infections ensue. This is fatal unless treated with antiviral drugs.

Once a virus has bound to its cellular receptor, the capsid penetrates the cell and its genome (DNA or RNA) is released into the cell cytoplasm. The main 'aim' of a virus is to reproduce successfully, and to do this its genetic material must download the information it carries. Mostly, this will take place in the cell's nucleus where the virus can access the molecules it needs to begin manufacturing its own proteins. Some large viruses, like pox viruses, carry genes for the enzymes they need to make their proteins and so are more self-sufficient and can complete the whole life cycle in the cytoplasm.

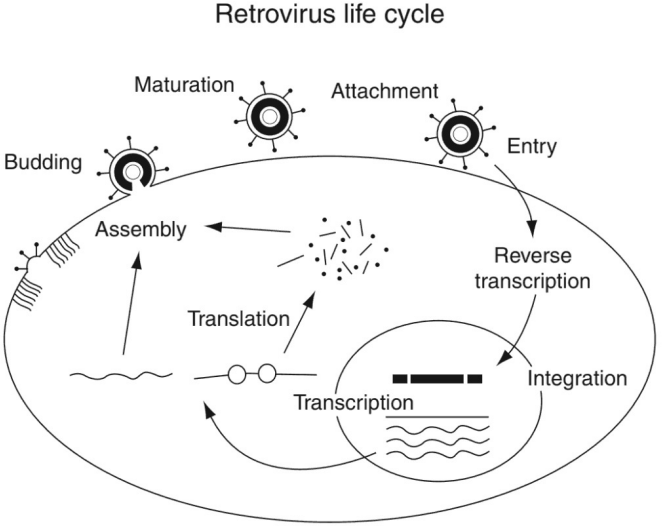
Once inside a cell, DNA viruses simply masquerade as pieces of cellular DNA, and their genes are transcribed and translated using as much of the cell's machinery as they require. The viral DNA code is transcribed into RNA messages which are read and translated into individual viral proteins by the cell's ribosomes. The separate virus components are then assembled into thousands

of new viruses which are often so tightly packed inside the cell that it eventually bursts open and releases them, inevitably killing the cell. Alternatively, new viruses leave rather more sedately by budding through the cell membrane. In the latter case, the cell may survive and act as a reservoir of infection.

RNA viruses are one step ahead of DNA viruses in already having their genetic code as RNA. As they carry enzymes that enable their RNA to be copied and translated into proteins, they are not so dependent on cellular enzymes and can often complete their life cycle in the cytoplasm without causing major disruption to the cell.

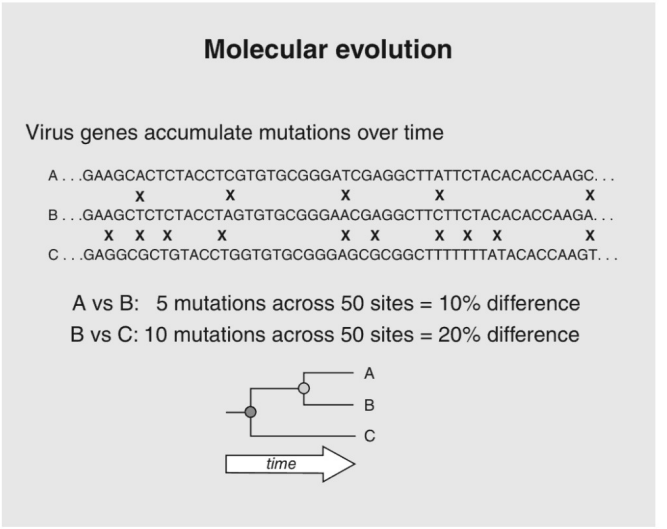
Retroviruses are a family of RNA viruses, including HIV, that have evolved a unique trick for establishing a lifelong infection of a cell while hiding from immune attack. Retrovirus particles contain an enzyme called *reverse transcriptase* which, once inside a cell, converts their RNA to DNA (Figure 4). This viral DNA can then join, or *integrate*, into the cell's DNA using another enzyme carried by the virus called *integrase*. The integrated viral sequence is called the *provirus*, and is effectively archived in the cell, remaining there permanently to be copied along with cellular DNA when the cell divides. The provirus is inherited by the two daughter cells, so building up a store of infected cells inside its host. At any time, a provirus can manufacture new viruses which bud from the cell surface, but in this instance it kills the cell.

In mammalian cells, the process of copying DNA during cell division is highly regulated, with a proof-reading system and several checkpoints in place to detect damaged or miscopied DNA and to correct the mistakes. If the damage is too great, cells have an 'auto-destruct' programme called *apoptosis* that induces death rather than allowing the cell to pass on its faulty DNA. Despite these checks, mistakes slip through, causing mutations to be replicated and passed on to future generations (Figure 5).



4. The retrovirus infectious cycle, showing viral entry into a cell followed by reverse transcription, integration, transcription and translation of the genome, virus assembly, and budding of new particles from the cell surface.

Virus genomes mutate far more quickly than the human genome, partly because viruses reproduce in a day or two with many thousands of offspring. Also, RNA viruses have no proof-reading system so they have a higher mutation rate than DNA viruses. Thus, every time a virus infects a cell, its DNA or RNA may be copied thousands of times, and as each new strand is incorporated into a new virus particle, every round of infection throws up several mutant viruses. This high mutation rate in viruses is their lifeline; in some, it is essential for their survival. Each round of infection produces some viruses that are non-viable due to mutations that interrupt the function of essential genes, and others with mutations that cause no change in function. But a few of the offspring will have beneficial mutations, giving them a selective advantage over their siblings. The benefit may result in



5. Molecular evolution of a viral gene over time. The extent of accumulated difference between the sequences is used to construct an evolutionary tree, in which the lengths of the horizontal branches are drawn to scale, and denote the time since common ancestors (shown as circles).

any number of advantages, including a heightened ability to hide from immune attack; to survive and spread between hosts; to resist antiviral drugs; or to reproduce at a faster rate. Whatever the advantage, it will lead to that particular mutant virus outstripping its siblings and eventually taking over in the population. Examples of this are common, particularly among RNA viruses like measles, which has been infecting the human population for at least 2,000 years. Despite this, scientists calculate that the present-day measles strain arose much more recently. Presumably, this virus strain was 'fitter' than its predecessor in some way; perhaps it had better spreading powers, and so eventually replaced the former strain worldwide. Another famous example is HIV, which rapidly evolves resistance to the drugs used to control the infection. In practice, this means

that several antiretroviral drugs have to be used together for effective treatment, and even then drug resistance is a growing problem. When a drug-resistant virus is transmitted to an uninfected person, the new infection is much more difficult to control. The same process has also foiled all attempts to make an effective HIV vaccine.

Analysing the mutations in its genome is a useful way of tracking a virus's history. The molecular clock hypothesis, which was developed in the 1960s, states that the mutation rate per generation is constant for any given gene. In other words, as applied to viruses, two samples of the same virus isolated at the same time from different sources will have evolved for the same length of time since their common ancestor. Since they will both have been accumulating mutations at a constant rate, the degree of difference between their gene sequences provides a measure of the time that has passed since their common ancestor. This way of measuring evolutionary time has been verified in higher life forms by comparing the dates of origin estimated by the molecular clock with those estimated from fossil records, but unfortunately viruses leave no such records. Nevertheless, scientists use the molecular clock to estimate the time of origin of certain viruses, and plot evolutionary (or phylogenetic) trees showing their degree of relatedness to other viruses. Because viruses have a high mutation rate, significant evolutionary change, estimated at around 1 per cent per year for HIV, can be measured over a short timescale. This technique was used to uncover the history of the measles virus. It was also used to discover that the smallpox virus is most closely related to the pox viruses of camels and gerbils, suggesting that all three arose from a common ancestor around 5,000 to 10,000 years ago.

Because virus particles are inert, without the ability to generate energy or manufacture proteins independently, they are not generally regarded as living organisms. Nonetheless, they are pieces of genetic material that parasitize cells, very efficiently