Van G. Wilson

Viruses: Intimate Invaders

palgrave
macmillan

Contents

1	**The Question of Life**	1
	Weird Facts and Big Numbers	1
	From Ignorance to Fascination	4
	A Biological Primer – Multicellular Organisms	6
	Smaller, But Still Alive – Unicellular Organisms	9
	Biochemistry – You Can't Escape It	11
	Viruses At Last	15
	Additional Reading	18
2	**Families, Forms, and Functions**	23
	Taxonomy and Other Geeky Things	23
	Virions, the Viral Vehicle	27
	Alive or Not, Viruses Have a Life Cycle	31
	Infection and Disease-How Viruses Spread and Do Nasty Things	36
	Additional Reading	42
3	**Ancient or New – On the Origin of Viruses**	47
	We Don't Know Exactly Where We Came From, But We're Here Anyway	47
	Are Viruses the Chicken or the Egg?	50
	Hypotheses, Hypotheses, and More Hypotheses	51
	New Data at Last	56
	Closer to Home	59
	Papillomaviruses and Hominid Evolution	63
	Additional Reading	65

4	**Of Predators and Prey**	69
	Viral Infections – The Good, the Bad, and the Ugly	69
	Influenza – More Than Just Bad Air	72
	Hepatitis C – The Great Deceiver	77
	Herpes Varicella-Zoster Virus – Now You See It, Now You Don't	81
	Polyomaviruses and Anelloviruses – Predators or Passengers?	85
	Additional Reading	89
5	**Immunity, Immunity, Immunity**	95
	The Big Three	95
	Intrinsic Immunity – Always There When We Need It	98
	Innate Immunity – Locked and Loaded	102
	Adaptive Immunity – The Gift That Keeps on Giving	109
	Additional Reading	116
6	**Viruses That Shaped Our World**	121
	Smallpox – Mankind's First Viral Conquest	122
	Polio – A Force for Change	129
	HIV/AIDS – A New Pandemic for the Twentieth Century and Beyond	137
	Additional Reading	146
7	**SV40 – An Obscure Monkey Virus and the Golden Age of Molecular Biology**	149
	Viruses, Viruses, and More Viruses	149
	Polio Vaccines and A Stealth Simian Virus	150
	To Be or Not To Be – Is SV40 a Human Tumor Virus?	153
	The SV40 Revolution	157
	p53 – The Guardian of the Genome	160
	Asilomar and the Birth of Biosafety	164
	Additional Reading	168
8	**Viral Cancer**	171
	The Infectious Beginnings of Cancer Research	171
	The Oncogene Revelation	174
	Human Retroviruses at Last	184
	Herpesviruses Turn Oncogenic	187
	Hepatitis Viruses and Hepatocellular Carcinoma	192
	Human Papillomaviruses – Cancer as an STD	195

Merkel Polyomavirus – Finally an Oncogenic Human Polyomavirus 197
The Seven Deadly Cancer Viruses 199
Additional Reading 200

9 The Viruses Within – How Human Are We? 203
Adam and EVEs 203
HERVs, HERVs, and More HERVs 205
Beyond HERVs 207
Assailants or Allies? 208
The HERVs That Made Us Human 210
The Dark Side 219
Additional Reading 223

10 Vaccines and the Conquest of Viruses 227
The Era of Vaccines 227
Vaccinology 101 233
The Road Is Long with Many a Winding Turn 236
Anti-Vaxxers: The Force Awakens 239
Vaccine Technology – A Glimpse of the Future 246
Additional Reading 251

11 New and Emerging Viruses 255
One Planet, Many Viruses, Countless Opportunities for Infections 255
The Unique Role of Bats 260
The Virus-Host Interface 262
Once and Future Threats: Influenza and Coronaviruses 267
Additional Reading 275

12 Beyond Antibiotics – Are Phages Our Allies? 279
From Fact to Fiction 279
A Brief History of Phage Therapy 280
A Phage Primer 283
The Antibiotic Problem 287
Phage Therapy Redux 292
A Glimpse of the Future 297
Additional Reading 299

13 Viruses in Gene Therapy and Cancer Therapeutics 303
Gene Therapy Emerges 303
From Adversary to Deliveryman – Harnessing Retroviruses for Gene Therapy 306
Into The Fray – What Other Viruses Bring to Gene Therapy 310
Oncolytic Viruses – Pitting One Killer Against Another 318
Additional Reading 324

Epilogue 329

Index 333

List of Figures

Fig. 1.1 **The composition and structure of DNA and RNA.** DNA is composed of four deoxyribonucleotides designated A, C, G, and T that are linked together in a linear strand. This linear sequence is called the primary structure. Our genomic DNA is double-stranded so two parallel chains bind to each other through the nucleotide pairs; A always pairs with T and C always pairs with G. Normally the double-stranded DNA is coiled into a helical form and this shape is referred to as the secondary structure. RNA is composed of similar nucleotides called ribonucleotides. For RNA there is no T ribonucleotide and instead, there is a ribonucleotide called U that pairs with A. RNA is typically single-stranded but can form short double-stranded regions through self-base pairing that creates structures called hairpin loops. These hairpin loops constitute the secondary structures of RNA. (Created with BioRender.com) 8

Fig. 1.2 **Protein composition and structure.** Proteins are comprised of 20 individual subunits called amino acids (3 are depicted). Amino acids are linked together into long chains that constitute the primary structure of proteins. Localized regions of these chains can fold in specific ways that are designated secondary structures. Finally, the entire chain will assume a specific three-dimensional shape that is referred to as the tertiary structure. Evolutionarily related proteins will exhibit similarities in the sequences and structure. (Created with BioRender.com) 12

Fig. 1.3 **DNA to Protein.** The so-called "central dogma" of biology is that genetic information flows from DNA to mRNA to proteins. DNA contains an organism's genes that are passed from generation to generation. To express a gene it is transcribed into a

complementary messenger RNA (mRNA) by the enzyme RNA polymerase. After its synthesis, the nucleotide sequence of the mRNA is read by ribosomes in increments of 3 nucleotides called codons. Transfer RNAs (tRNAs) carry amino acids to the ribosome and each tRNA recognizes a specific codon. The ribosome helps position the tRNAs on the mRNA and then catalyzes the joining of adjacent amino acids to form the protein chain. The finished protein is released from the ribosome to perform its function in the cell. (Created with BioRender.com) 14

Fig. 2.1 **Linnaean classification of a representative Virus Realm.** The figure depicts the Realm of Monodnaviria and its descending subdivisions from Kingdoms to Species. This Realm has four Kingdoms, and each Kingdom has the subdivision shown below the Kingdom. The number of further subdivisions within each classification is as indicated. For example, the Kingdom Loebivirae has 1 Phylum, 1 Class, 1 Order, 2 Families, 0 Subfamilies, 26 Genera, and a total of 34 Species 25

Fig. 2.2 **Depiction of the genome types.** Shown are the four possible types of DNA or RNA genomes: (1) double-stranded, circular; (2) single-stranded, circular; (3) double-stranded, linear; and (4) single-stranded, linear. Types 1–4 have been found in DNA viruses while only types 2–4 have been observed in RNA viruses. (Created with BioRender.com) 26

Fig. 2.3 **Morphology of Poxviruses.** Shown is a poxvirus virion with the structural features labeled. The poxvirus core is the nucleocapsid consisting of a capsid protein structure containing the viral DNA genome and various other viral proteins. The core has a figure-eight shape with protein structures called lateral bodies located in the indentations of the core. The entire core is surrounded by a membrane envelope embedded with viral proteins. (Created with BioRender.com) 29

Fig. 2.4 **The viral life cycle.** The seven steps of the viral life cycle from cell attachment (1) to release (7) are illustrated with the entry step (2) depicted in more detail in Fig. 2.5. The life cycle shown would be typical of an RNA virus since it is occurring in the cytoplasm, while DNA viruses would be performing their gene expression (4) and genome replication (5) steps in the cell nucleus. (Created with BioRender.com) 31

Fig. 2.5 **Viral entry mechanisms.** Nonenveloped viruses enter cells through the invagination of the cell membrane in a process known as endocytosis. Internalized virions are temporarily contained in intracellular vesicles and must disrupt the vesicle to release the virion into the cytoplasm. Some enveloped viruses are

also taken up by endocytosis while others enter via a membrane fusion process. In the latter process, the virion membrane fuses with the cell membrane which releases the nucleocapsid into the cytoplasm. (Created with BioRender.com) 32

Fig. 3.1 **The tree of life.** On primordial Earth, there was likely a multitude of primitive replicative entities (proto-viral RNAs, protocells, and/or proto-virocells) that were the precursors to modern cells and viruses. Included in this mix were mobile genetic elements (MGEs) that were small segments of RNA that could jump from one location in a larger RNA to another location within the same RNA. From this complex and competing mix entities arose LUCA, the Last Universal Common Ancestor. LUCA became the dominant life form and is the seed of the Tree of Life that gave rise to all three of the modern lineages of life, bacteria, eukaryotes, and the archaea. Because all life on Earth arose from a single common ancestor, all life shares a common genetic code and uses the same basic process to convert DNA to RNA to protein. (Created with BioRender.com) 49

Fig. 3.2 **The tree of modern virus lineages.** From an ancestral pre-virus emerged the initial positive-sense RNA (PS-RNA) virus whose genomic RNA could be directly translated into proteins by cellular ribosomes. The primordial PS-RNA virus eventually gave rise to three PS-RNA superfamilies including the picornavirus superfamily, the alphavirus superfamily, and the flavivirus superfamily. Each superfamily gave rise to several of the modern PS-RNA virus families. Additionally, the negative-sense RNA (NS-RNA) viruses arose from the flavivirus superfamily. NS-RNA viruses have a genome that cannot be directly translated by ribosomes, and instead, the NS-RNA must first be converted to PS-RNA after infection. The primordial PS-RNA virus also gave rise to most of the double-stranded RNA (DS-RNA) viruses except for Reoviruses that have a different origin unrelated to the PS-RNA viruses. In addition to the DS-RNA viruses, the PS-RNA viruses also gave rise to the small single-stranded DNA (SS-DNA) viruses that became the polyomaviruses and the papillomaviruses. The larger DNA viruses have distinct lineages with herpesviruses forming one evolutionary group and the giant viruses, poxviruses, and adenoviruses forming a second distinct evolutionary group. (Created with BioRender.com) 61

Fig. 4.1 **Viral reassortment.** Shown are two related influenza viruses, Parental Virus A and Parental Virus B. Both parental virions contain the 8 different RNA segments that constitute the complete influenza virus genome (blue RNA segments for virion A

and red RNA segments for virion B). The surface proteins of these two parental viruses are antigenically different as represented by the different colors for these proteins. During a mixed infection, the 8 genome segments from each parental virus will be replicated, and the cells will fill with a mixture of many red and blue copies of each segment. As new virions assemble, each progeny virion must package one copy of RNA segments 1–8. For each of these 8 RNA segments packaged there is a random choice between a blue copy and a red copy. Some progeny virions may contain all blue copies or all red copies identical to the parental viruses. However, most virions likely possess a mixture of some red and some blue copies and these are the reassortants. The reassortant shown has 6 blue RNA segments derived from Parental Virus A and 2 red segments derived from Parental Virus B, but every combination of red and blue mixed reassortants is possible. In the example shown, the mixed RNA segments in the reassortant encode surface proteins from each parent virus resulting in a reassortant that is antigenically different from either parent virus. Such changes in reassortants can lead to novel viruses with properties very different from the parental viruses. (Created with BioRender.com) 74

Fig. 4.2 **Varicella-Zoster virus latency.** The figure depicts a single neuron innervating a region of the skin. The cell nucleus is located within the cell body, and a single long projection called an axon stretches from the cell body to the tissue being innervated. Cell bodies from many individual neurons group together in structures near the spinal cord called ganglia (not shown). When varicella-zoster virus (HVZ) infects the skin it replicates productively in the epithelial cell where it produces chickenpox lesions. Some virions may enter the axon where their viral genomic DNA is transported within the axon to the cell body nucleus. The HVZ DNA remains in the cell nucleus in an inactive state without viral reproduction. This is a permanent condition that persists for the lifetime of the infected person. As immunity wanes with aging, viral DNA in one or a few cells may reactivate and travel back down the axon to the original region of skin that the axon innervated. Upon reaching the skin the virus will start the productive replication cycle that produces new virions and causes the skin eruption of shingles. The location where shingles occurs on the body is random and depends on which neuronal cell suffered the reactivation event. (Created with BioRender.com) 83

Fig. 5.1 **Innate immunity.** Depicted are the recognition and signaling steps of the innate response by a cell infected with a DNA virus.

After attachment and entry of the viral particle, virion uncoating releases viral DNA into the cytoplasm. The cytoplasm is normally free of DNA, so the viral DNA is a pathogen-associated molecular pattern (PAMP) that warns the cell of a viral infection. The viral DNA is bound by cytoplasmic receptors (pattern recognition receptors – PRRs) that specifically recognize DNA. The binding of viral DNA to the PRR initiates a signaling cascade that transmits to the cell nucleus to turn on the production of various proteins. Among the proteins produced in response to viral infection are cytokines and interferon that are released from the infected cell as danger signals to warn surrounding uninfected cells that a virus is present. (Created with BioRender.com) 103

Fig. 5.2 **Interferon signaling.** Interferon release by an infected cell can diffuse to nearby cells or enter the bloodstream and become more systemic. Cells with an interferon receptor can bind the circulating interferon which activates a signal cascade in the receiving cell. The signal generated by interferon binding to its receptor transmits a message to the cell nucleus to induce the expression of dozens to hundreds of genes that are collectively referred to as interferon-stimulated genes (ISGs). Among the ISG protein products are the viral restriction factors (VRFs) and protein kinase R (PKR) which acts as an antiviral factor. (Created with BioRender.com) 104

Fig. 5.3 **Protein kinase R.** Protein kinase R (PKR) is one of several anti-viral proteins induced by interferon. In response to interferon, cells begin to make larger amounts of PKR, but the enzyme exists in an inactive form in the cytoplasm of the producing cells. For activation, PKR requires long, double-stranded RNA (dsRNA), a molecule not usually available in the cytoplasm. However, if the cell becomes infected, particularly with RNA viruses, then long, ds RNA becomes abundant as the viral RNA genomes copy themselves as part of the replication process. When ds RNA is present it binds to PKR and converts PKR from an inactive protein to an active enzyme. Active PKR now modifies a second protein that is critical for cellular protein synthesis, resulting in the inhibition of protein synthesis. Without protein synthesis, the cell cannot produce new virions, thus helping to limit the spread of the virus. (Created with BioRender.com) 108

Fig. 5.4 **T cell activation.** Activation of a resting T cell requires interaction with an antigen-presenting cell (APC) that displays a foreign antigen (for example a viral protein fragment) bound to its major histocompatibility complex II (MHC II). For the interaction to occur, the T cell receptor must recognize the foreign antigen

associated with the APC MHC II. The formation of this cell-to-cell complex via antigen recognition triggers the T cell to activate and stimulates its proliferation. (Created with BioRender.com) 111

Fig. 5.5 **B cell activation.** Resting B cells express both B cell receptors (BCRs) and major histocompatibility locus II (MHC II) proteins on their surface. The BCR is a membrane-bound form of the specific antibody that a B cell can produce. The binding of a virion surface protein to the B cell receptor triggers step 1 of the B cell activation process. Bound virions are internalized by the B cell, degraded into protein fragments, and a fragment recognized by the BCR is returned to the B cell surface for display with the MHC II. Step 2 of the activation requires a helper T cell with a T cell receptor (TCR) that can recognize the protein fragment displayed by the B cell. The binding of the T cell to the B cell via the MHC II-fragment-TCR linkage activates the resting B cell and converts it into an antibody-secreting plasma cell. The antibody produced is directed against whatever viral protein was responsible for the initial interaction between the virus and the BCR. (Created with BioRender.com) 114

Fig. 8.1 **The retroviral life cycle.** Retroviruses have a single-stranded RNA genome and their virions are unusual in that each virion contains two copies of the genomic RNA. In addition to the RNA, the virion also carries the viral reverse transcriptase (RT) enzyme. After virion binding to the cell receptor (step 1) and fusion of the viral and cell membranes (step 2), the viral RNA is released into the cytoplasm along with the RT. The RT uses the genomic RNA as a template to produce a double-stranded DNA copy (step 3) while simultaneously degrading the RNA. The double-stranded DNA copy is transported to the nucleus and randomly integrates into a host chromosome (step 4) to generate the proviral form. Once integrated, the provirus is transcribed into RNA using the host cell RNA polymerase (step 5). The newly transcribed RNAs are both translated into new viral protein (step 6) and used as new genomic RNAs for packaging into progeny virions (step 7). The progeny virions bud out of the cell and acquire a membrane envelope from the cell membrane (step 8) and are ultimately released (step 9) to infect other cells. (Created with BioRender.com) 176

Fig. 8.2 **Organization of retroviral genomes.** Depicted are the proviral forms that consist of double-stranded DNA encoding the viral genes flanked on both ends with the long terminal repeat (LTR) sequences. The LTRs contain the transcriptional regulatory elements (promoters and enhancers) that control the transcription

of the viral genes. Leukemia viruses are wild-type viruses fully capable of replication and reproduction to generate progeny virions. These viral genomes contain the four basic retroviral genes: Gag (capsid proteins), Pro (a protease enzyme), Pol (the reverse transcriptase and integrase), and Env (surface proteins). Sarcoma viruses typically have lost some of the standard four genes and have replaced the missing viral sequences with an oncogene derived from a host proto-oncogene. Because they lack a complete complement of the four essential leukemia virus genes, sarcoma viruses are defective and are unable to reproduce by themselves. Instead, sarcoma viruses only reproduce during a co-infection with a complementary leukemia virus that supplies the missing protein products. The trans-acting viruses are a subfamily of retroviruses known as lentiviruses. The hallmark of lentiviruses is the presence of other genes (indicated in the pX region) in addition to the four standard leukemia virus genes. Shown is the genome organization for a representative lentivirus, HTLV-1. Tax and Hbz are the genes whose protein products are implicated in the transforming activity of this virus. The arrow above the Hbz gene denotes that it is transcribed from the right-end LTR in contrast to all the other viral genes that are transcribed from the left-end LTR. (Created with BioRender.com) 178

Fig. 8.3 **The promoter insertion model of Leukemia virus transformation.** When leukemia viruses infect cells they randomly integrate the double-stranded DNA form of their genomes into host chromosomes. When the integration occurs in a harmless region of the chromosomal DNA then there is no damage from the proviral integration and the cells remain normal. However, when integration occurs proximal to a proto-oncogene then the proto-oncogene can be constitutively transcribed into mRNA by the adjacent viral long terminal repeat (LTR). The continuous translation of the proto-oncogene mRNA into the protein product can drive the transformation of this cell. (Created with BioRender.com) 183

Fig. 10.1 **Types of vaccines**. Viral vaccines are categorized as inactivated vaccines, attenuated vaccines, or subunit vaccines. Inactivated vaccines use whole virions and chemically treat the virions to render them noninfectious. The virions still retain their morphology and antigenicity which allows them to induce neutralizing antibodies. Since the inactivated vaccine virus is "killed" it cannot cause the original viral disease. Attenuated vaccines use a mutated version of the pathogen that has greatly reduced virulence. These weakened viruses still infect cells and replicate to a limited degree

which generally induces a robust and long-lived B and T cell immune response. However, attenuated vaccines may still cause mild disease in normal recipients and should not be used in immunocompromised patients or pregnant women in whom they might cause serious disease. Subunit vaccines use only a portion of the virion, typically one or more surface proteins that will invoke neutralizing antibodies against the whole virion. The vaccine can consist of purified viral protein or can be the gene (DNA) or mRNA that encodes the target protein. The pathogen gene is often administered by cloning it into a harmless virus to make a chimeric delivery virus while mRNA is more commonly delivered via lipid nanoparticles. Subunit vaccines contain neither the whole virion nor the complete viral genome, therefore, they cannot cause the viral disease. (Created with BioRender.com) 234

Fig 11.1 **Receptors for avian influenza virus.** Influenza viruses use their attachment protein, the viral hemagglutinin (HA) protein, to bind to receptors on host cells. Sialic acid (SA) in the alpha-2,3 form (SA-alpha-2,3) is the receptor that interacts with the avian influenza HA protein, while human influenza viruses use SA-alpha-2,6 as their receptor. The alpha-2,3 form of SA is found on the cells in the digestive tract of avian species, the lower respiratory tract of humans, and the upper respiratory tract of swine. Cells in the upper respiratory tract of humans have the SA-alpha-2,6 form of sialic acid and this receptor is not suitable for avian influenza viruses. However, cells with the SA-alpha-2,3 form that are present in the lower human respiratory tract can be infected by avian influenza if the virus reaches this region. Swine have both the SA-alpha-2,3 and the SA-alpha-2,6 forms on cells in their upper respiratory tract so they are susceptible to both avian and human influenza viruses. (Created with BioRender.com) 264

Fig. 12.1 **Phage structure.** The figure shows two phages, one phage just after attachment to the surface of a bacterium (**a**) and a second phage injecting its genome into the bacterium (**b**). Depicted is a tailed phage that consists of two parts, a head structure (the virus capsid equivalent) containing the phage genome and a complex tail structure composed of a collar, a sheath, a base plate, and fibers as shown. The tail fibers are the appendages that interact with receptor molecules on the surface of a bacterium to initiate the binding of the phage. After the phage-receptor interaction forms, the tail contracts which drives the sheath through the bacterial membrane and cell wall. Penetration creates a pathway through the sheath for the phage nucleic acid to leave the head

and enter the cytoplasm of the bacterium where it can begin to replicate. (Created with BioRender.com) 284

Fig. 12.2 **Mechanisms of resistance transfer.** (**a**) Conjugation is a mating process between two bacteria where a physical connection forms between the two cells. The connection allows nucleic acids to pass from one bacterium (the donor) to a recipient bacterium. In the example shown the upper bacterium is the donor. This donor bacterium contains its DNA chromosome as well as small, circular DNA plasmids carrying an antibiotic resistance gene (orange segment). Transfer of one or more plasmid copies to the recipient cell will endow the recipient bacterium with antibiotic resistance. (**b**) As bacteria die and lyse, their DNA fragments (both the plasmids and the chromosome) are released into the environment. Transformation is the process where bacteria randomly take up these DNA fragments in the environment and can incorporate those fragments into their genomes. If a bacterium incorporates a DNA fragment containing an antibiotic resistance gene (orange fragment) then that bacterium transforms into an antibiotic-resistant strain. (**c**) Transduction is the process of bacteria acquiring new genes when they are infected by a lysogenic phage. During infection, the phage genome will enter the bacterium and be integrated into the bacterial genome. If that phage carries an antibiotic resistance gene (orange segment) then that gene will become part of the bacterial genome and will confer resistance on the bacterium. Collectively these three gene transfer mechanisms enable bacteria to rapidly acquire and swap antibiotic resistance. (Created with BioRender.com) 289

Fig. 13.1 **Adenovirus vectors.** Shown are the wild-type adenovirus genome and three generations of vectors derived from the wild-type virus. The arrows indicate the positions of the early genes E1-E4. The Δ symbol indicates a deleted gene and the Ψ symbol shows the location of the sequence needed for DNA packaging into adenovirus virions. The ITRs are the inverted terminal repeat sequences present at each end of the adenovirus genome. These ITRs are the only sequences needed for adenovirus-directed replication of DNA. Any exogenous therapeutic gene being delivered by the vector is the transgene. (Created with BioRender.com) 312

Fig. 13.2 **Oncolytic adenoviruses.** Wild-type adenoviruses express both the E1B-55K and E1A proteins (E1B-55K$^+$ and E1A$^+$). These viral proteins allow the wild-type adenovirus to overcome the cell-protective effects of the p53 and pRB proteins in normal cells, resulting in cell death (solid arrow). Wild-type adenoviruses also effectively kill tumor cells (solid arrow) as these cells frequently

lack p53 ($p53^-$) or have bypassed the pRB/E2F checkpoint. Because wild-type adenoviruses not only kill tumor cells but also normal cells in the patient, they are not suitable for tumor therapy. In contrast, the oncolytic adenoviruses do not produce a functional version of either the E1B-55K protein (the ONYX-015 virus) or the E1A protein ($E1A^-$ viruses). Viruses lacking E1B-55K or E1A are defective for replication in normal cells because of the virus-inhibitory activities of cellular p53 and pRB. Consequently, the oncolytic adenoviruses do not kill normal cells effectively (dashed arrow). However, tumor cells often lack p53 ($p53^-$) or have bypassed the pRB/E2F checkpoint. Without these protective mechanisms, the tumor cells become susceptible to infection and killing by the oncolytic adenoviruses (solid arrow). This selective killing of tumor cells and not normal cells makes oncolytic adenoviruses potentially effective for therapeutic application against appropriate cancers. (Created with BioRender.com) 319