"...everywhere an interplay between nucleic acids and proteins; a spinning wheel in which the thread makes the spindle and the spindle the thread"

--ERWIN Chargaff
Virology breakthrough in the 1950’s:

The viral nucleic acid genome is the genetic code

Hershey-Chase experiment with phage T4

Fraenkel-Conrat’s work with TMV
Alfred Hershey & Martha Chase, 1952

Viral protein labeled with radioactive sulfur

Infection

Radioactivity predominantly in the supernatant fraction

Centrifugation/detection

No radioactivity detected in next generation of phage

Viral DNA labeled with radioactive phosphorus

Blending/separation

Radioactivity predominantly in the cell pellet

Centrifugation/detection

Radioactive DNA is detected in progeny phage
The bigger surprise: thousands of different virions, seemingly infinite complexity of infections

*But a finite number of viral genomes*
Key fact makes your life easier:

Viral genomes must make mRNA that can be read by host ribosomes

All viruses on the planet follow this rule, no exception to date
David Baltimore (Nobel laureate) used this insight to describe a simple way to think about virus genomes.

The original Baltimore system missed one genome type: the gapped DNA of the Hepadnaviridae.
Definitions

- mRNA (ribosome ready) is always the plus (+) strand
- DNA of equivalent polarity is also the (+) strand
- RNA and DNA complements of (+) strands are negative (-) strands
- Not all (+) RNA is mRNA!
The elegance of the Baltimore system

Knowing only the nature of the viral genome, one can deduce the basic steps that must take place to produce mRNA.
The seven classes of viral genomes

- dsDNA
- gapped dsDNA
- ssDNA
- dsRNA
- ss (+) RNA
- ss (-) RNA
- ss (+) RNA with DNA intermediate
Go to:

b.socrative.com/login/student
room number: virus

Why is mRNA placed at the center of the Baltimore scheme?

A. Because all virus particles contain mRNA  
B. There is no specific reason  
C. Because all viral genomes are mRNAs  
D. Because mRNA must be made from all viral genomes  
E. Because Baltimore studied mRNA
Viral DNA or RNA genomes are structurally diverse

- Linear
- Circular
- Segmented
- Gapped
- Single-stranded (+) strand
- Single-stranded (-) strand
- Single-stranded, ambisense
- Double-stranded
- Covalently attached proteins
- Cross-linked ends of double-stranded DNA
- DNA with covalently attached RNA
What is the function of genome diversity?

- DNA and RNA based
  - RNA genomes appeared first in evolution (RNA World)
  - Switch to DNA genomes
  - Only RNA genomes on planet today are viral
  - Viroids: Relics of RNA world?
- Linear, circular, segmented, ds, ss, (+), (-)
Memorize 7 genome types and key virus families

If you know the genome structure you should be able to deduce:
- How mRNA is made from the genome
- How the genome is copied to make more genomes
What information is encoded in a viral genome?

*Gene products and regulatory signals for:*

- Replication of the viral genome
- Assembly and packaging of the genome
- Regulation and timing of the replication cycle
- Modulation of host defenses
- Spread to other cells and hosts
Information NOT contained in viral genomes

- No genes encoding the complete protein synthesis machinery (AARS, elFs, tRNAs)
- No genes encoding proteins involved in energy production or membrane biosynthesis
- No classical centromeres or telomeres found in standard host chromosomes
- Probably we haven’t found them yet - 90% of giant virus genes are novel
# Largest known viral genomes

<table>
<thead>
<tr>
<th>Virus</th>
<th>Length</th>
<th>Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pandoravirus salinus</td>
<td>2,473,870</td>
<td>2,541</td>
</tr>
<tr>
<td>Pandoravirus dulcis</td>
<td>1,908,524</td>
<td>1,487</td>
</tr>
<tr>
<td>Megavirus chilensis</td>
<td>1,259,197</td>
<td>1,120</td>
</tr>
<tr>
<td>Mamavirus</td>
<td>1,191,693</td>
<td>1,023</td>
</tr>
<tr>
<td>Mimivirus</td>
<td>1,181,549</td>
<td>979</td>
</tr>
<tr>
<td>Moumouivirus</td>
<td>1,021,348</td>
<td>894</td>
</tr>
<tr>
<td>Mimivirus M4</td>
<td>981,813</td>
<td>620</td>
</tr>
<tr>
<td>C. roenbergenensis virus</td>
<td>617,453</td>
<td>544</td>
</tr>
<tr>
<td>Mollivirus sibericum</td>
<td>651,000</td>
<td>523</td>
</tr>
<tr>
<td>Pithovirus sibericum</td>
<td>610,033</td>
<td>467</td>
</tr>
</tbody>
</table>
## Smallest known viral genomes

<table>
<thead>
<tr>
<th>Virus</th>
<th>Length</th>
<th>Protein</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viroid</td>
<td>120</td>
<td>none</td>
</tr>
<tr>
<td>Satellite</td>
<td>220</td>
<td>none</td>
</tr>
<tr>
<td>Hepatitis delta satellite</td>
<td>1,700</td>
<td>1</td>
</tr>
<tr>
<td>Circovirus</td>
<td>1,759</td>
<td>2</td>
</tr>
<tr>
<td>Anellovirus</td>
<td>2,170</td>
<td>4</td>
</tr>
<tr>
<td>Geminivirus</td>
<td>2,500</td>
<td>4</td>
</tr>
<tr>
<td>Hepatitis B virus</td>
<td>3,200</td>
<td>7</td>
</tr>
<tr>
<td>Levivirus</td>
<td>3,400</td>
<td>4</td>
</tr>
<tr>
<td>Partitivirus</td>
<td>3,700</td>
<td>2</td>
</tr>
<tr>
<td>Barnavirus</td>
<td>4,000</td>
<td>7</td>
</tr>
</tbody>
</table>
Go to:

b.socrative.com/login/student
room number: virus

What information may be encoded in a viral genome?

A. Gene products that catalyze membrane biosynthesis
B. Gene products that catalyze energy production
C. Complete protein synthesis systems
D. Centromeres or telomeres
E. Enzymes to replicate the viral genome
Viral DNA genomes

- The host genetic system is based on DNA
- Many DNA viruses emulate the host
- However, almost all viral DNA genomes are NOT like cell chromosomes
- Unexpected tricks have evolved
dsDNA genomes

Adenoviridae

Herpesviridae

Papillomaviridae

Polyomaviridae

Poxviridae
dsDNA genomes

Genomes copied by host DNA polymerase

Polyomaviridae (5 kbp)

Papillomaviridae (8 kbp)

Genomes encode DNA polymerase

Adenoviridae (36–48 kbp)

Herpesviridae (120–220 kbp)

Poxviridae (130–375 kbp)
Gapped dsDNA genomes

*This genome cannot be copied to mRNA*

Hepadnaviridae
Hepatitis B virus
ssDNA genomes

Circoviridae (1.7–2.2 kb)

TT virus (ubiquitous human virus)

Parvoviridae (4–6 kb)

B19 parvovirus (fifth disease)
Which DNA genome, on entry into the cell, can be immediately copied into mRNA?

A. dsDNA
B. gapped dsDNA
C. circular ssDNA
D. linear ssDNA
E. All of the above
RNA genomes

- Cells have no RNA-dependent RNA polymerase (RdRp)
- RNA virus genomes encode RdRp
- RdRp produce RNA genomes and mRNA from RNA templates
dsRNA genome

Reoviridae (19–32 kbp in 10 dsRNA segments)

Rotavirus (human gastroenteritis)
ssRNA: (+) sense

*Picornaviridae* (Poliovirus, Rhinovirus)

*Caliciviridae* (gastroenteritis)

*Coronaviridae* (SARS)

*Flaviviridae* (Yellow fever virus, West Nile virus, Hepatitis C virus, Zika virus)

*Togaviridae* (Rubella virus, Equine encephalitis virus)
ssRNA: (+) sense

*Coronaviridae* (28–33 kb)

5' UTR → A₇A₂₅OH 3'

*Flaviviridae* (10–12 kb)

5' UTR → 3'

*Picornaviridae* (7–8.5 kb)

5' VPg UTR → A₇A₂₅OH 3''

*Togaviridae* (10–13 kb)

5' UTR → A₇A₂₅OH 3'

**Genome**

**– RNA**

© Principles of Virology, ASM Press
ssRNA(+) sense with DNA intermediate

One viral family: *Retroviridae*

Two human pathogens:

Human immunodeficiency virus (HIV)
Human T-lymphotropic virus (HTLV)
The remarkable retroviral genome strategy

Retroviridae (7–10 kb)

proivirus

\[ +RNA \rightarrow -DNA \rightarrow DNA \rightarrow +RNA \]

© Principles of Virology, ASM Press
ssRNA, (-) sense

*Paramyxoviridae* (Measles virus, Mumps virus)

*Rhabdoviridae* (Rabies virus)

*Filoviridae* (Ebola virus, Marburg virus)

*Orthomyxoviridae* (Influenza virus)

*Arenaviridae* (Lassa virus)
ssRNA, (-) sense

Segmented genomes: Orthomyxoviridae (10–15 kb in 6–8 RNAs)
(-) strand RNA segments

Nonsegmented genomes: Paramyxoviridae (15–16 kb)

Rhabdoviridae (13–16 kb)
Reassortment: Consequence of segmented genome

L

M

L

M

R3

L R3 M

L M
Ambisense RNA genomes

Genome 5' c 3'
Antigenome 3' c 5'

Replication
Transcription
Transcription

GP mRNA 5' c

NP mRNA 5' c

Arenaviridae
RNA pol in virion
A

Linear (+) strand RNA genome of a picornavirus

5' VPg UTR

A_{r\alpha} A_{OH} 3'

B
Go to:

b.socrative.com/login/student
room number: virus

Which statement about viral RNA genomes is correct?

A. (+)ssRNA genomes may be translated to make viral protein
B. dsRNA genomes can be directly translated to make viral protein
C. (+)ssRNA virus replication cycles do not require a (-) strand intermediate
D. RNA genomes can be copied by host cell RNA-dependent RNA polymerases
E. All of the above
This method allowed the application of genetic methods to animal viruses
Engineering mutations into viral genomes - the modern way

- Infectious DNA clone: transfection
- A modern validation of the Hershey-Chase experiment (1952)
- Deletion, insertion, substitution, nonsense, missense
Genetic methods

**Transfection**

- Production of infectious virus after transformation of cells by viral DNA, first done with bacteriophage lambda

- Transformation-infection
**Fig. 03.12**

**A**

Poliovirus

(+) Viral RNA

Infection

Transfection

Cultured cells

In vitro RNA synthesis

Poliovirus DNA

cDNA synthesis and cloning

**B**

Poliovirus DNA

In vitro RNA synthesis

(+) strand RNA transcript
Resurrecting the 1918 influenza virus

- Influenza virus was not identified until 1933
- In 2005, influenza RNA was isolated from formalin-fixed, paraffin-embedded lung tissue sample from autopsy of victim of influenza in 1918
- Influenza RNA also isolated from frozen sample obtained by in situ biopsy of the lung of a victim buried in permafrost since 1918
- Complete nucleotide sequence of all 8 RNA segments determined
- Virus was recovered by transfection of cells with 8 plasmids containing genome sequences
Synthetic Virology and Biosecurity

NSABB: National Science Advisory Board for Biosecurity

- Federal advisory committee to provide advice, guidance, and leadership regarding biosecurity oversight of *dual use research* to all Federal departments and agencies with an interest in life sciences research.

- Advises on and recommends specific strategies for the efficient and effective oversight of federally conducted or supported dual use biological research, taking into consideration national security concerns and the needs of the research community.