

Classification and Taxonomy

Viruses are highly diverse infectious agents that differ in their genome composition, morphology, replication strategies, and host range. To understand this diversity, virologists use classification systems that organize viruses according to their biological characteristics and evolutionary relationships.

Two major approaches are used in modern virology:

1. ICTV classification: A formal taxonomic system that organizes viruses into hierarchical groups, primarily according to their evolutionary relationships and genomic characteristics.
2. Baltimore classification: A functional system that divides viruses into seven groups according to their genome type and the mechanism used to produce messenger RNA (mRNA).

2. ICTV Classification of Viruses

2.1. Definition and role of the ICTV

The International Committee on Taxonomy of Viruses (ICTV) is the official international body responsible for developing, maintaining, and updating a universal system for virus taxonomy and nomenclature.

It operates within the Virology Division of the International Union of Microbiological Societies (IUMS).

The main objectives of the ICTV are to:

- Establish an internationally recognized classification system for viruses.
- Define and standardize the names of viral taxa.
- Organize viruses according to their evolutionary relationships.
- Review and approve changes in virus taxonomy.
- Maintain an official database of recognized viral taxa.

The ICTV system is continuously updated as new viruses are discovered and additional genomic and evolutionary evidence becomes available. The official taxonomy is published in the Master Species List (MSL). The 2025–2026 release, MSL41, recognizes 10 realms, 427 families, 4,149 genera, and 17,554 species.

2.2. Principles of virus classification

Virus classification is based on several biological, molecular, and evolutionary characteristics.

A. Genome characteristics

The nature and organization of the viral genome are fundamental classification criteria. These include:

- Nucleic acid type: DNA or RNA.
- Strandedness: single-stranded (ss) or double-stranded (ds).
- Genome structure: linear or circular.

- Genome organization: segmented or non-segmented.
- Genome size and gene arrangement.

B. Genomic sequence similarity and phylogenetic relationships

Comparative genomics allows researchers to identify conserved genes, determine genetic similarities, and reconstruct evolutionary relationships among viruses.

Phylogenetic analysis is particularly important for defining taxa and determining whether newly identified viruses belong to existing groups.

C. Virion morphology

Structural characteristics include:

- Capsid symmetry, such as icosahedral or helical organization.
- Presence or absence of a lipid envelope.
- Virion shape, size, and structural complexity.

D. Replication and gene-expression strategies

Viruses may differ in their mechanisms of genome replication, transcription, translation, and viral protein synthesis.

E. Host range and biological properties

Host specificity, tissue tropism, pathogenicity, antigenicity, and vector specificity may also contribute to classification, particularly when distinguishing closely related species.

2.3. ICTV taxonomic hierarchy

The ICTV uses a hierarchical system comparable to the taxonomic organization of cellular organisms. It uses characteristic suffixes to identify taxonomic ranks.

Taxonomic rank	Standard suffix
Realm	-viria
Subrealm	-vira
Kingdom	-virae
Subkingdom	-virites
Phylum	-viricota
Subphylum	-viricotina
Class	-viricetes
Subclass	-viricetidae
Order	-virales
Suborder	-virineae

Family	-viridae
Subfamily	-virinae
Genus	-virus
Subgenus	-virus
Species	Binomial name; no fixed suffix

These suffixes facilitate the recognition of taxonomic ranks and ensure consistency in virus nomenclature.

2.5. Virus names versus species names

A virus is a biological entity, whereas a virus species is a taxonomic category.

Under current ICTV nomenclature rules:

- Species names consist of two components: a genus name followed by a species epithet.
- Official taxon names are written in italics.
- Ordinary virus names are generally written in non-italic type.

3. Baltimore Classification of Viruses

3.1. General principles

The Baltimore classification was introduced by David Baltimore in 1971 as a functional approach to understanding viral genome expression. Unlike ICTV taxonomy, which emphasizes evolutionary relationships, Baltimore classification groups viruses according to two fundamental characteristics:

1. The nature of their genetic material.
2. The pathway used to generate messenger RNA (mRNA).

This system originally included six groups. A seventh group was subsequently added to accommodate DNA viruses that replicate through reverse transcription.

3.2. The seven Baltimore groups

Group	Genome type	Pathway to mRNA	Examples
I	dsDNA	DNA → mRNA	Herpesviruses, adenoviruses
II	ssDNA	ssDNA → dsDNA → mRNA	Parvoviruses
III	dsRNA	dsRNA → mRNA	Rotaviruses
IV	(+)ssRNA	Genome functions as mRNA	Poliovirus, SARS-CoV-2
V	(-)ssRNA	(-)RNA → mRNA	Influenza virus, rabies virus
VI	ssRNA-RT	RNA → DNA → mRNA	HIV-1
VII	dsDNA-RT	DNA → RNA → DNA during replication	Hepatitis B virus

RT = reverse transcription.

3.3. Enzymes involved in viral genome expression

Several polymerases participate in the transcription and replication of viral genomes.

Abbreviation	Enzyme	Main function
DdRP	DNA-dependent RNA polymerase	Produces RNA from a DNA template
DdDP	DNA-dependent DNA polymerase	Synthesizes DNA from a DNA template
RdRP	RNA-dependent RNA polymerase	Synthesizes RNA from an RNA template
RT	Reverse transcriptase (RNA-dependent DNA polymerase activity)	Produces DNA from an RNA template

4. Replication Strategies of the Seven Baltimore Groups

4.1. Group I: Double-stranded DNA (dsDNA) viruses

Group I viruses contain double-stranded DNA genomes, which may be linear or circular depending on the virus. Their genome organization resembles cellular DNA in the sense that DNA serves as a template for both transcription and genome replication.

Pathway to mRNA: **dsDNA** → **mRNA** → **Viral proteins**

Viral mRNA is synthesized from the DNA genome by a DNA-dependent RNA polymerase (DdRP). Depending on the virus, transcription is performed by the host cell's RNA polymerase or by a virus-encoded RNA polymerase. The resulting mRNAs are translated into viral proteins by host ribosomes.

The viral DNA genome is replicated by a DNA-dependent DNA polymerase (DdDP).

Some DNA viruses rely extensively on cellular replication machinery, whereas others encode their own DNA polymerases and accessory proteins.

Examples: Herpes simplex virus, human adenoviruses, human papillomaviruses, polyomaviruses, poxviruses, and numerous bacteriophages.

4.2. Group II: Single-stranded DNA (ssDNA) viruses

Group II viruses contain single-stranded DNA genomes. The genomic DNA may be linear or circular. Unlike dsDNA viruses, these viruses must generally synthesize a complementary DNA strand before transcription can occur.

Pathway to mRNA: **ssDNA** → **dsDNA intermediate** → **mRNA** → **Viral proteins**

After entering a suitable host cell, the ssDNA genome is converted into a double-stranded DNA replicative form. This intermediate serves as the template for the synthesis of viral mRNAs by a DNA-dependent RNA polymerase.

The mRNAs are subsequently translated by host ribosomes.

New viral DNA molecules are produced from the dsDNA replicative intermediate. Depending on the virus family, replication may involve rolling-circle or rolling-hairpin mechanisms.

Examples: Parvoviruses, circoviruses, geminiviruses, and genomoviruses.

4.3. Group III: Double-stranded RNA (dsRNA) viruses

Group III viruses possess double-stranded RNA genomes, which can be segmented or non-segmented. The RNA duplex contains complementary positive-sense and negative-sense strands.

Pathway to mRNA: **dsRNA → (+)mRNA → Viral proteins**

RNA-dependent RNA polymerase is required

Double-stranded RNA cannot be translated directly by host ribosomes. A virus-encoded RNA-dependent RNA polymerase (RdRP) synthesizes positive-sense mRNA using the negative-sense strand as a template. These positive-sense transcripts are then translated by host ribosomes to produce viral proteins.

Host cells do not normally possess the RdRP activity required to transcribe viral dsRNA into mRNA. Consequently, dsRNA viruses generally package the necessary polymerase in their infectious particles.

Examples: Rotaviruses, orthoreoviruses, birnaviruses, and partitiviruses.

4.4. Group IV: Positive-sense single-stranded RNA [(+)ssRNA] viruses

Group IV viruses contain single-stranded RNA genomes with the same polarity as messenger RNA. Therefore, the genomic RNA can function as mRNA after its release into the appropriate cellular compartment.

Pathway to mRNA and protein synthesis: **Genome (+)ssRNA → Direct translation → Viral proteins**

Unlike negative-sense RNA viruses, positive-sense RNA viruses can generally initiate translation without first transcribing their genomes. The incoming RNA is translated by host ribosomes to produce viral proteins, including proteins involved in RNA replication.

After the viral replication proteins have been synthesized, a virus-encoded RdRP produces a complementary negative-sense RNA intermediate. This intermediate serves as the template for the synthesis of new positive-sense RNA molecules.

Examples: Poliovirus, hepatitis C virus, dengue virus, and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2).

4.5. Group V: Negative-sense single-stranded RNA [(-)ssRNA] viruses

Group V viruses contain single-stranded RNA genomes whose polarity is complementary to that of mRNA. As a result, their genomic RNA cannot be directly translated by host ribosomes.

Pathway to mRNA: **(-)ssRNA → (+)mRNA → Viral proteins**

Viral RdRP is required for transcription

The virus must first produce positive-sense mRNA from its negative-sense genomic RNA. This transcription is carried out by a virus-encoded RdRP. The resulting mRNAs are translated by the host cell to synthesize viral proteins.

Most animal negative-sense RNA viruses replicate in the cytoplasm. However, influenza viruses and bornaviruses are notable exceptions because they carry out important replication and transcription steps in the nucleus. Examples: Influenza viruses, rabies virus, measles virus, Ebola virus, and numerous plant viruses.

4.6. Group VI: Single-stranded RNA viruses with reverse transcription (ssRNA-RT)

Group VI viruses contain positive-sense single-stranded RNA [(+)ssRNA], but they use an enzyme called reverse transcriptase (RT) to convert their RNA into DNA. These viruses are mainly called retroviruses. Example: Human immunodeficiency virus (HIV).

Replication steps (HIV):

1. Entry: The virus enters the host cell and releases its RNA genome.
2. Reverse transcription: Reverse transcriptase converts the viral RNA into double-stranded DNA.
3. Integration: The viral DNA is inserted into the host cell's DNA by an enzyme called integrase, forming a provirus.
4. Transcription: The provirus is used to produce viral mRNA and new genomic RNA.
5. Protein synthesis: Host ribosomes translate viral mRNA into viral proteins.
6. Assembly: Viral RNA and proteins are assembled into new viral particles.

Difference between Groups IV and VI:

- Group IV: The positive-sense RNA genome can be translated directly into viral proteins.
- Group VI: The viral RNA is first converted into DNA, which is then used to produce viral mRNA and new genomic RNA.

Other example: Rous sarcoma virus.

4.7. Group VII: Double-stranded DNA viruses with reverse transcription (dsDNA-RT)

Group VII viruses contain DNA, but they use an RNA intermediate to produce new viral DNA genomes. This process requires an enzyme called reverse transcriptase (RT).

Example: Hepatitis B virus (HBV).

Replication steps (HBV):

1. Entry: The virus enters the host cell and releases its partially double-stranded DNA.
2. DNA repair: In the nucleus, viral DNA is converted into a complete circular DNA molecule called cccDNA.
3. Transcription: The cccDNA is used to produce viral mRNA and pregenomic RNA (pgRNA).
4. Protein synthesis: Host ribosomes translate viral mRNA into proteins.
5. Reverse transcription: The pgRNA is converted into new viral DNA by reverse transcriptase inside the viral capsid.

Difference between Groups I and VII:

- Group I: DNA is replicated directly from DNA.
- Group VII: DNA is produced from an RNA intermediate by reverse transcription.

Other example: Cauliflower mosaic virus (a plant virus).