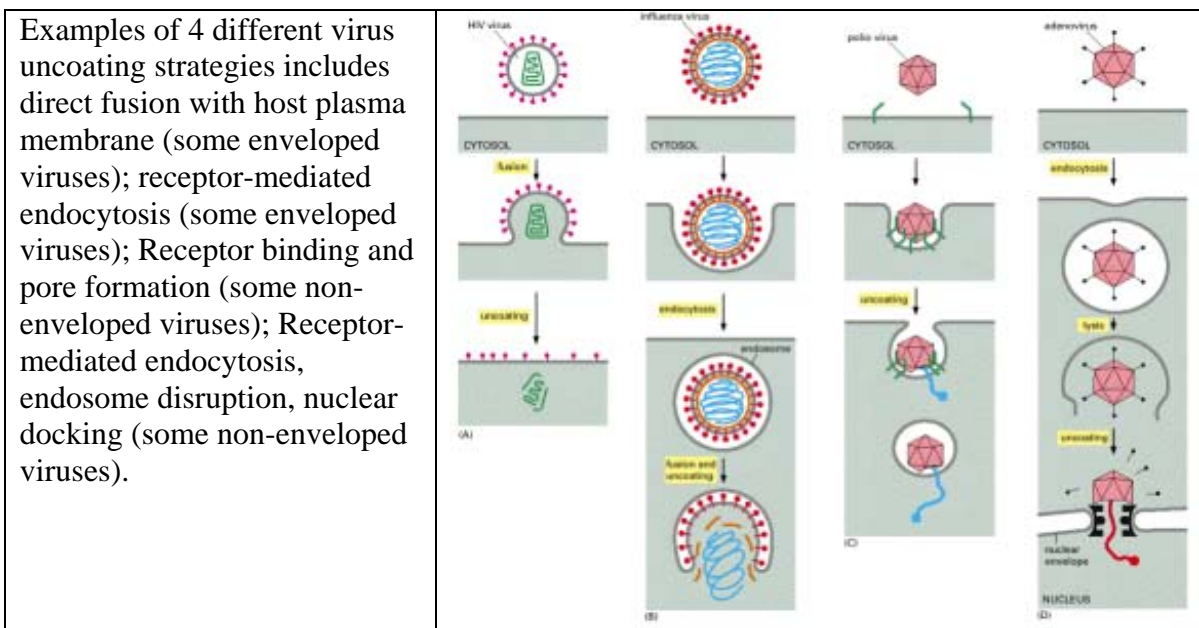# Viruses and Gene Expression

As we know viruses are parasites, therefore they can only reproduce by invading a controlling other cells as they lack the cellular machinery for self-reproduction. They are composed of a genetic code, which may be DNA or RNA, and it may be single or double stranded. The tail contains apparatus for injecting the DNA into a host. This head is a protein coat that surrounds the virus, and it is called the Capsid. The capsid can appear in many different shapes.

Viral replication cycle makes use of host's gene regulation machinery. Viruses must enter host cells and uncoat.

Examples of 4 different virus uncoating strategies includes direct fusion with host plasma membrane (some enveloped viruses); receptor-mediated endocytosis (some enveloped viruses); Receptor binding and pore formation (some non-enveloped viruses); Receptor-mediated endocytosis, endosome disruption, nuclear docking (some non-enveloped viruses).	

DNA viruses tend to have genes with powerful gene regulatory sequences. For example in the experiment shown in class the endogenous plant had a promoter driving the expression of the reporter gene GUS, and we can observe very little and localized expression patterns. However, when the promoter of the virus (Cauliflower Mosaic Virus – CaMV 35S) was used in the plant we could observe a massive expression of the reporter GUS, on virtually every cell in that plant.

|	
RNA viruses on the other hand need to replicate their genome & express their proteins. A good example of this would be the Poplar Mosaic Virus Genome. Its genome is composed of different parts shown on the picture below.	

Smith & Campbell (2004) Archives of Virology
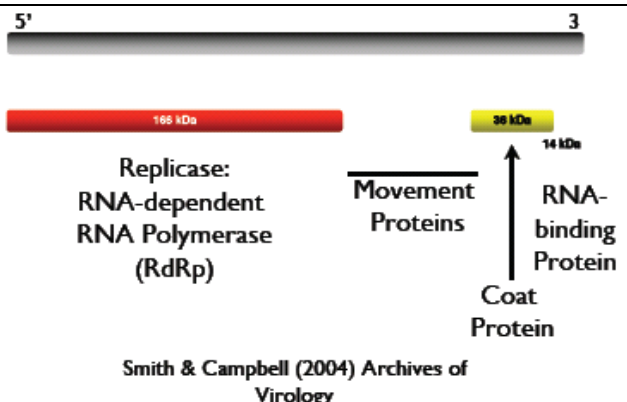
Now, some single-stranded RNA viruses replicate their genome using an RNA-dependent RNA polymerase (RdRp). The process starts with:
- The RNA virus infiltrating the plant cell.
- RdRp synthesizes new RNA strand to from a template.
- RdRp synthesizes new RNA starand in 5' to 3' direction.
- In plants, movement proteins enable the viral genome to move from cell to cell through the plasmodesmata.

A virus generally enters a new cell through an insect vector, unless the transmission is within the organism. After entering the new cell it is then assembled. There are many types of viruses so you can imagine that there are several ways a virus can be assembled. A spherical virus for example is a coated virus. This coat is formed by multiple protein subunits packed together to create the capsid. The capsid is self-assembled as shown in fig 3.32. The tomato bushy stunt virus is an example of a spherical virus. 180 identical coat proteins make up its capsid and it contains the 4500 nt RNA genome which is inside.

There are some viruses however, that are enveloped by a lipid bilayer derived from the host plasma membrane. A term to know is budding, which is involved in the envelope assembly. This lipid bilayer from the host is also combined with transmembrane viral envelope proteins that are encoded in the virus genetic code.

Bacteriophage lambda is a "virus" that attacks bacterias. This virus has two different replication pathways. The lysogenic cycle, which is when the viral genome is integrated to the host's genome, this is done without the requirement of ATP or DNA ligase (see details*). The lytic cycle, occurs when the virus replicates and leaves the cell when found under stress. These cycles are well presented on the picture below.
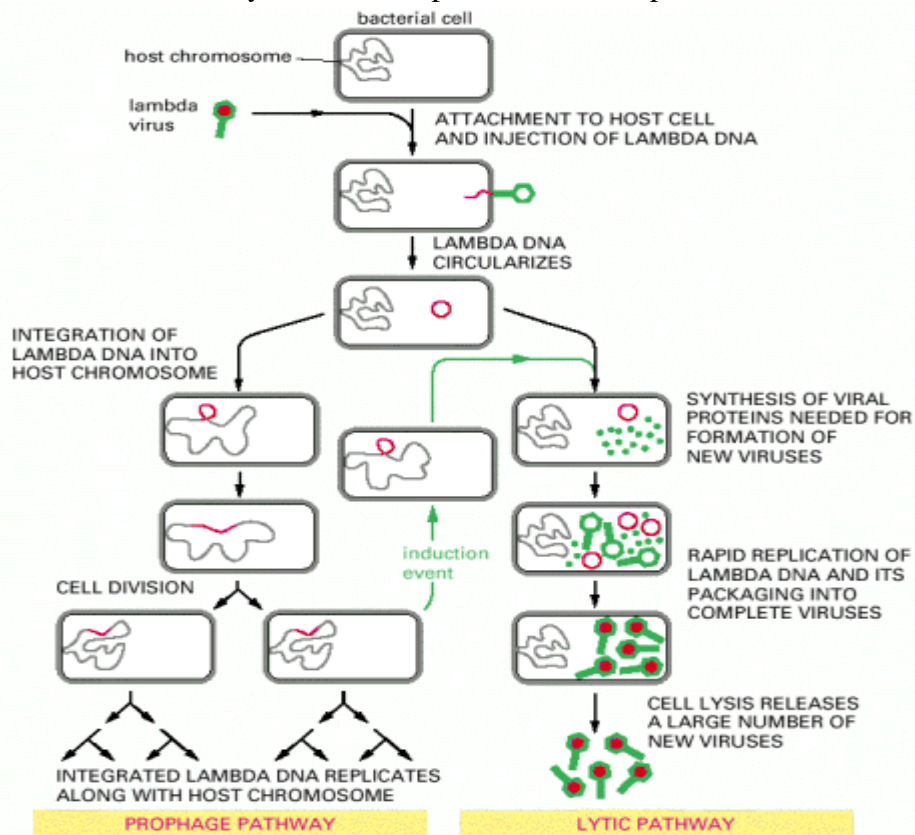
The way the Lambda virus is able to insert itself on the hosts genome is with the assistance of some proteins, and if we remember the Helix-turn-helix DNA binding motifs, we will remember that the four examples given were the trp repressor, CAP fragment, lambda Cro and lambda repressor fragment. The last two are the ones that control the bacteriophage lambda gene expression. Now on the picture below is a simplified representation of the regulatory system used by these bateriophage lambda in *E. coli*. In the stable state 1, during the lysogenic cycle, a repressor protein is synthesized, which activates its own synthesis and turns off the synthesis of several other phage proteins. Now during the lytic cycle, the phage synthesizes the Cro protein, which then turns off the synthesis of the repressor protein, so that many bateriophage proteins can be made and the viral DNA can replicate freely in the host cell. This eventually produces too many new phages, which ends up killing the cell.
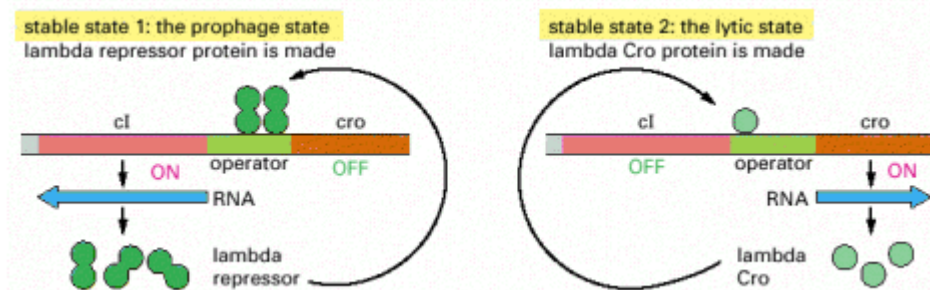
On the picture we can see two different genes encoding for two proteins. In the first one, the cI promoter, the distance between the operator and the promoter allows lambda repressor to promote RNA polymerase binding. So the repressor in this instance actually assists transcription. Now on the gene 2, Cro promoter, the distance between the operator and the promoter causes the lambda repressor to prevent RNA polymerase binding. So the same lambda repressor can assist or prevent RNA polymerase binding. For more info click here.

Now let's look at the retrovirus replication cycle. The retrovirus consists of two RNA molecules packed into each viral particle, the enzyme reverse transcriptase, the caspid and the envelope. Upon the viral entry on the host the reverse transcriptase first makes a DNA copy of the viral RNA molecule and then a second DNA strand of that sequence, generating a double stranded DNA copy of the RNA genome. Then the integrase enzyme encoded by the virus, catalyses the integration this DNA double helix into the host chromosome. The integration of the viral genome is required for the synthesis of new viral RNA molecules by the hosts RNA polymerase. Just note that the reverse transcriptase is composed of a polymerase domain covalently attached to RNAse Hybrid domain. This allows the protein to degrade the RNA strand when the RNA/DNA helix is going through, which helps the later synthesis of the second helix. An illustration of the whole process is on the next page.
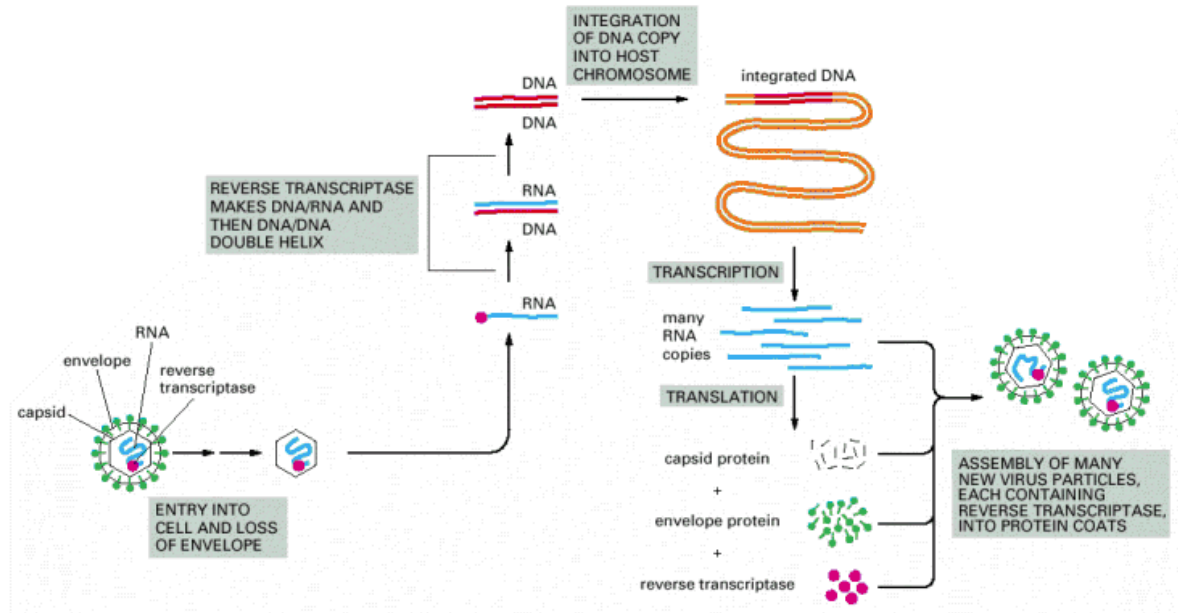
Host Defence

Now what kinds of defenses are available for the hosts?
- Sometimes none.
- Adaptive Immune Responses.
- Innate Immune Responses.
- Virus-induced gene silencing / RNA interference hosts defence mechanism.

Very recent studies have investigated a **Virus-induced gene silencing (VIGS)** mechanism, which can be used as a defense mechanism. Let's see how this works:
- It starts with RNA virus invading a cell, in this example we will look at a plant cell.
- As we have seen earlier, RdRp synthesizes a new RNA strand which is then made into a double RNA strand.
- Then a "surveillance system" recognizes and degrades this dsRNA into 21-23 nt fragments. The DICER, which is an RNAase, is the specific protein that does this surveillance.
- Then a multi-protein complex called **RISC** (RNA-induced silencing complex) binds to these dsRNA and tries to find more complementary strands to this sequence. Once these homologous sequences are found, they are "silenced".
- This "silencing" effect can spread throughout the plant, meaning it can go from cell to cell.
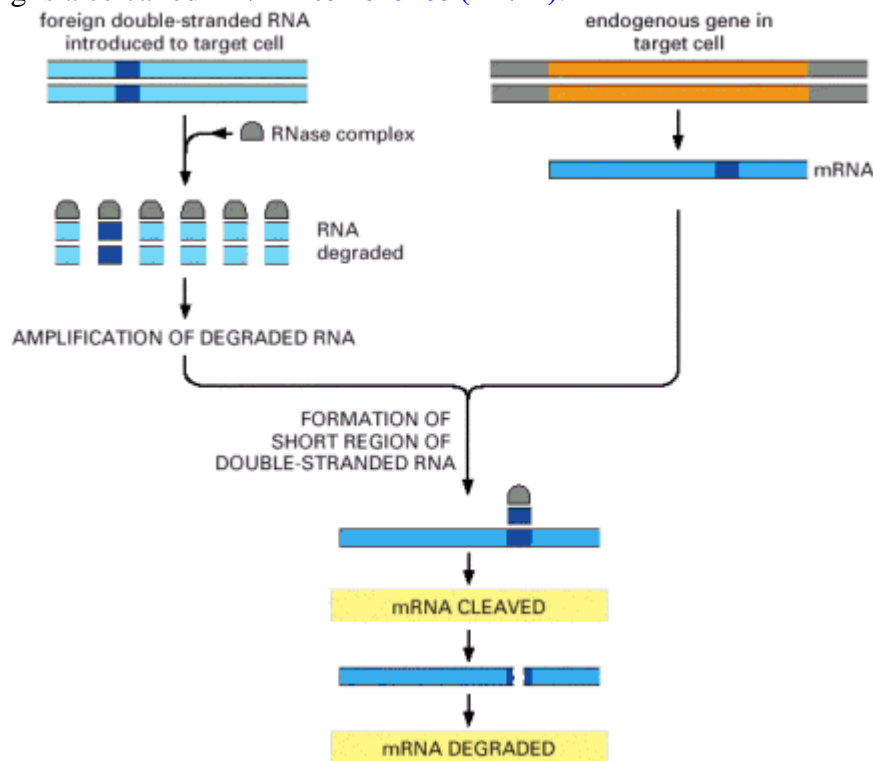
VIGS experiments have been done successfully in plants:
- An mRNA is synthesized from the host's genome.
- This mRNA is injected into a viral genome.
- This virus containing a segment of the host's RNA then infects the plant.
- RISC "silences" RNA molecules with sequence homology to the fragments.

- However, this sequence is homologous to the hosts DNA, therefore silencing of the hosts gene also occurs in the hosts genome.
- VIGS of host's gene is thought to occur via DNA methylation-mediated chromatin remodeling. Recall DNA methylation locks genes in a silent state, preventing genes from being expressed. Also this is epigenetic, which means it only happens to the organism in which it occurs, it does not pass on to future generations.

Evidence for VIGS has been shown with the RNA genome virus Tobacco Rattle Virus (TRV). TRV was genetically engineered to contain a piece of the plant RNA encoding Phytoene Desaturase (PDS), which is an enzyme that is essential to maintain chlorophyll pigments of the plant. Infection of the plants with the genetically engineered TRV containing PDS resulted in the silencing of the host PDS gene resulting in "bleached" plants lacking chlorophyll pigment.

This VIGS has also been shown to happen in animals. The dsRNA-mediated gene silencing is also called **RNA interference (RNAi)**.

foreign double-stranded RNA introduced to target cell

endogenous gene in target cell

RNase complex

RNA degraded

mRNA

AMPLIFICATION OF DEGRADED RNA

FORMATION OF SHORT REGION OF DOUBLE-STRANDED RNA

mRNA CLEAVED

mRNA DEGRADED

Another process, which relates to RNAi is the Antisense RNA, which is actually thought to work via RNAi. Basically, you have an organism with a certain gene X. Now if you want to silence that gene, you could theoretically add a dsRNA on the cell, which would work the way we described above. However, there is a more effective way to do this. What you would do is to insert on the genome a mutated gene X, and this mutated gene encodes for the antisense RNA of the normal gene X. The antisense RNA is simply a complementary strand of the sense (normal) RNA of the gene. So if this antisense RNA is transcribed in a significant quantity it will eventually inactivate the gene.

RNAi has proven to be a powerful experimental tool. Being used in E.Coli expressing double stranded RNA, which had been eaten by worm (*C.elegans*).

The RNAi technique has been widely used to study gene function in the nematode *C. elegans*. When working with worms, introducing the dsRNA is quite simple: RNA can be injected directly into the intestine of the animal, or the worm can be fed with *E. coli* expressing the target gene dsRNA (Figure 8-66A). The RNA is distributed throughout the body of the worm and is found to inhibit expression of the target gene in different tissue types. Further, as explained in Figure 7-107, the interference is frequently inherited by the progeny of the injected animal. Because the entire genome of *C. elegans* has been sequenced, RNAi is being used to help in assigning functions to the entire complement of worm genes. In one study, researchers were able to inhibit 96% of the approximately 2300 predicted genes on *C. elegans* chromosome III. In this way, they identified 133 genes involved in cell division in *C. elegans* embryos (Figure 8-66C). Of these, only 11 had been previously ascribed a function by direct experimentation.
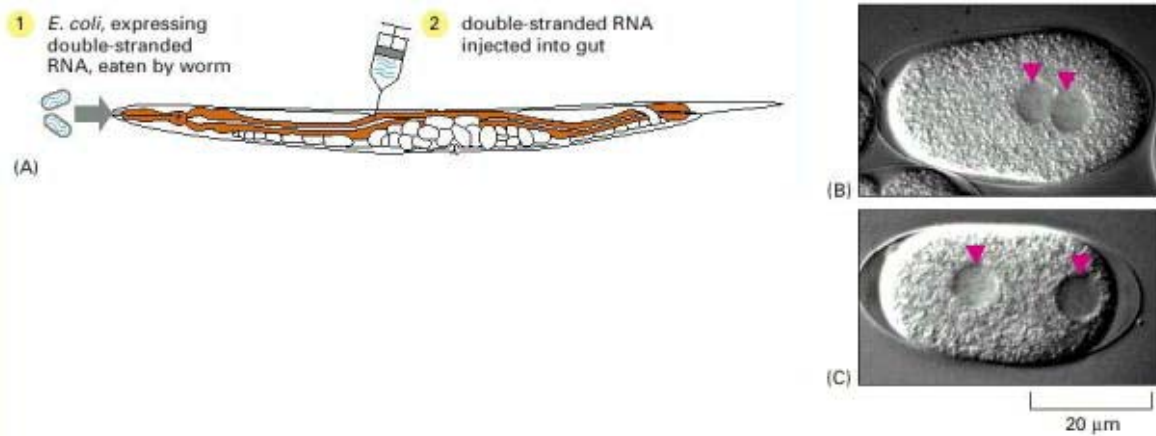

Figure 8-66. Dominant negative mutations created by RNA interference. (A) Double-stranded RNA (dsRNA) can be introduced into *C. elegans* (1) by feeding the worms with *E. coli* expressing the dsRNA or (2) by injecting dsRNA directly into the gut. (B) Wild-type worm embryo. (C) Worm embryo in which a gene involved in cell division has been inactivated by RNAi. The embryo shows abnormal migration of the two unfused nuclei of the egg and sperm. (B, C, from P. Gönczy et al., *Nature* 408:331–336, 2000. © Macmillan Magazines Ltd.)