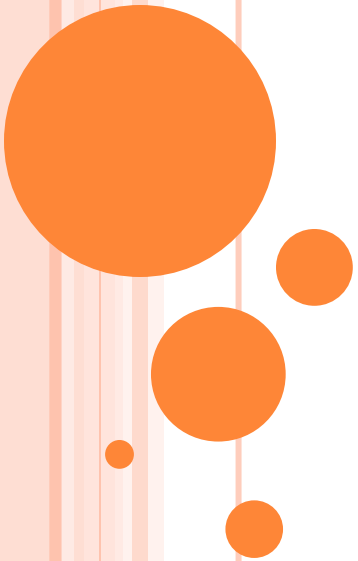
SALIENT FEATURES OF VIRAL GENOMES



unusual bases

other bases in addition to the normal adenine, cytosine, guanine and uracil. Found primarily in tRNAs and produced by the post trancription modification of one of the normal bases.

1) Unusual bases (TMV, T4 Phage)
2) Overlapping genes (hepatitis B virus)
3) Alternate splicing (retrovirus)
4) Terminal redundancy (T4 phage)
5) Terminal cohesive ends (lambda phage
6) *AMBISENSE GENOME*
7) *PARTIAL DOUBLE STRANDED GENOME*
8) *LONG TERMINAL REPEATS*
9) *SEGMENTED AND NON SEGMENTED GENOMES*
10) *CAPPING AND TAILING*

- Viruses have genome with certain unusual bases.
- Replacement of thymine by uracil in DNA bacteriophages. (T4)
- Thymine is also replaced by hydroxy methyl uracil.
- In certain phages, cytosine is replaced by 5-hydroxy methyl cytosine to protect phage DNA from host restriction
- and hydroxy methyl group is Glucosylated to protect phage DNA from host Mcr (modified cytosine restriction (T4).

- Definition of overlapping genes

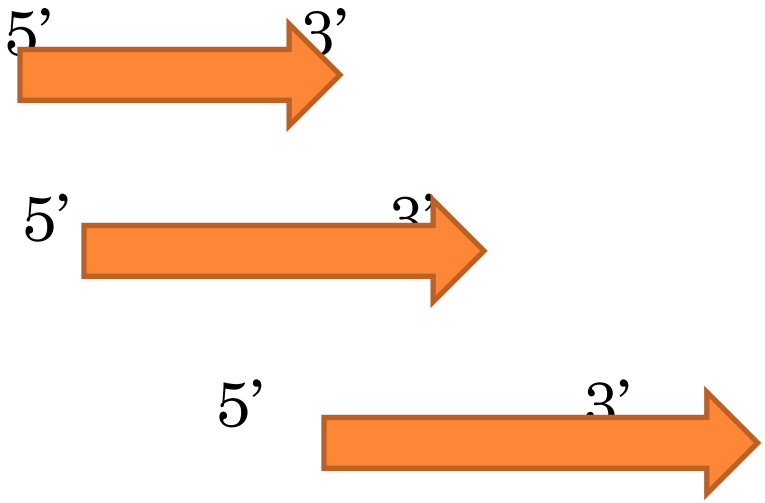
- Overlapping genes are defined as a pair of adjacent genes whose coding regions are partially overlapping. In other words, a single stretch of DNA codes for portions of two separate proteins. Such an arrangement of genetic code is ubiquitous. Many overlapping genes have been identified in the genomes of **prokaryotes**, **eukaryotes**, **mitochondria**, and viruses.

- ORFs in different reading frames may overlap each other.

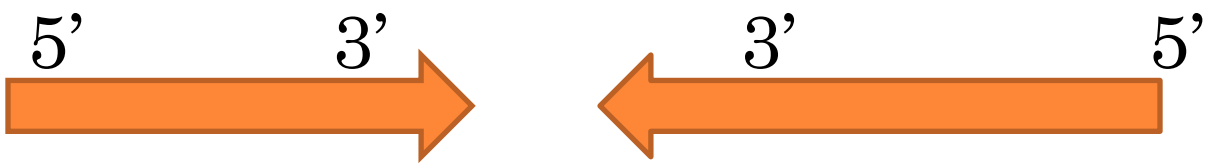
These are classified into 3 categories:

1) Unidirectional

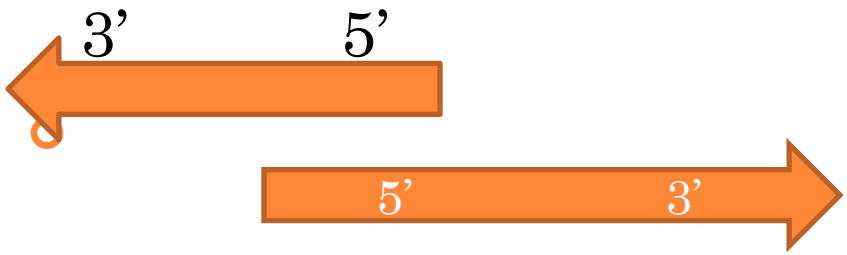
It is found most commonly



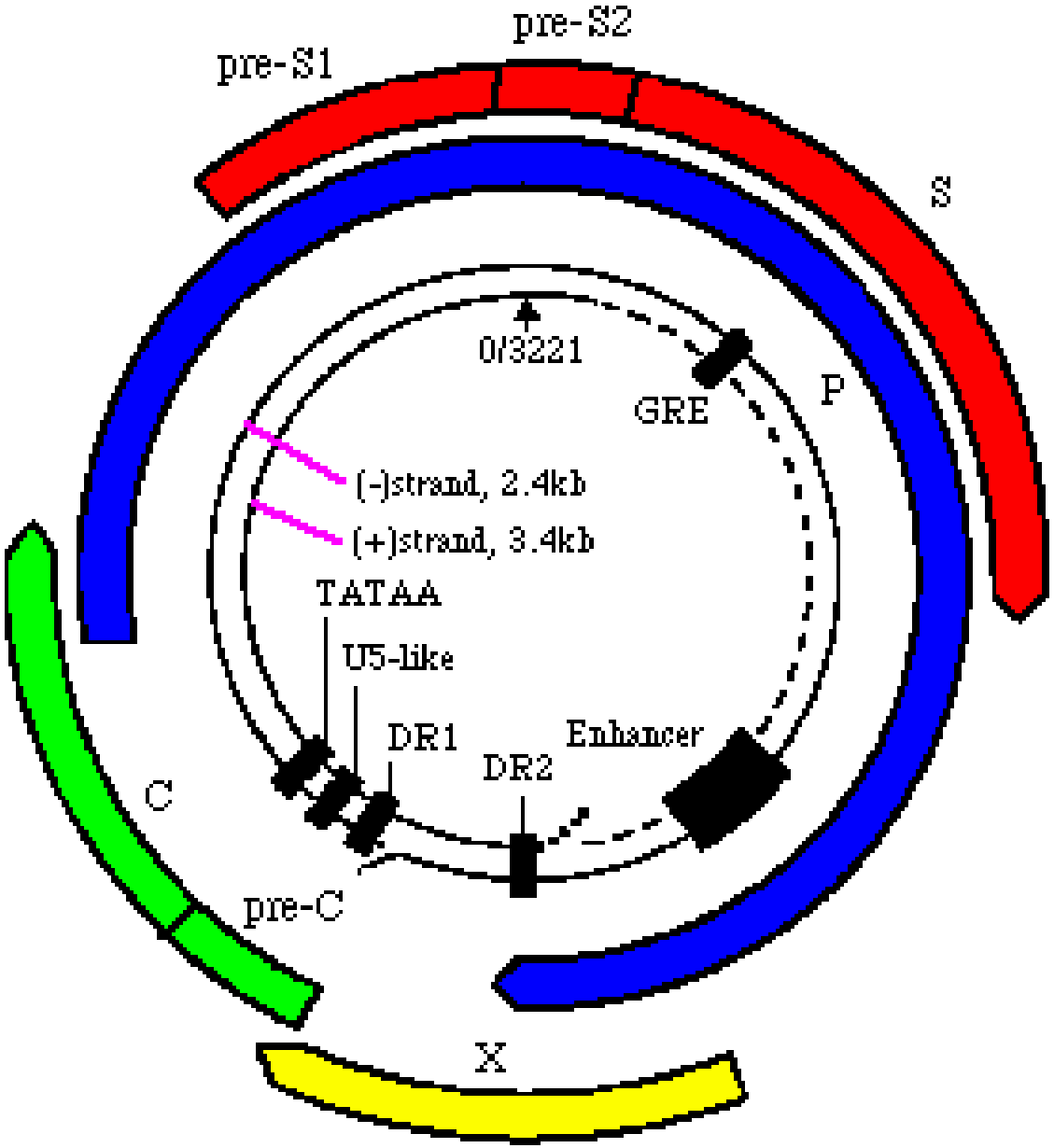
- 2) convergent



- 3)divergent

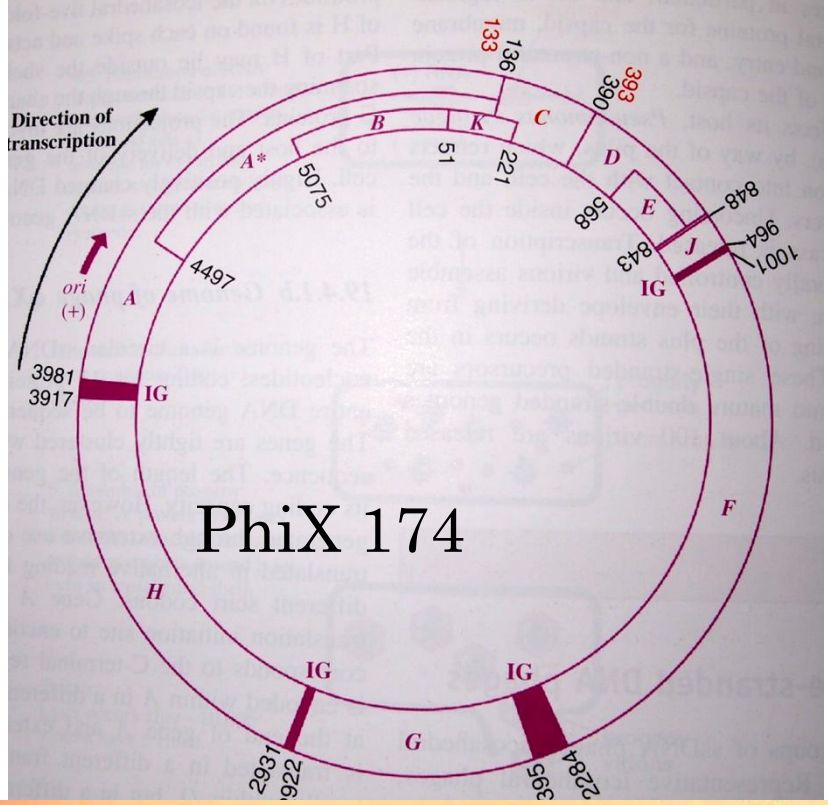


- Example : ø×174



S gene-surface GP, HBSag
P gene-comprises 80% of the genome encodes a polymerase : can act as a DNA pol., RT, and RNase H.

- The C gene encodes the HBcAg (core antigen).
- The X gene encodes essential proein and is thought to be associated with transcriptional activation.



gene *A* (3981–136) protein, viral strand synthesis and RF replication;
gene *A** (4497–136) protein, shutting down host DNA synthesis;
gene *B* (5075–51) protein, capsid morphogenesis;
gene *C* (133–393) protein, DNA maturation;
gene *D* (390–848) protein, capsid morphogenesis;
gene *E* (568–843) protein, host cell lysis;
gene *F* (1001–2284) protein, capsid morphogenesis – major coat protein;
gene *G* (2395–2922) protein, capsid morphogenesis – major spike protein;
gene *H* (2931–3917) protein, capsid morphogenesis – minor spike protein;
gene *J* (848–964) protein, capsid morphogenesis – core protein (DNA condensing protein);
gene *K* (51–221) protein, function not clear-appears to enhance phage yield (burst size).
IG: intergenic region at borders of genes *A, J, F, G, H* contains a ribosome binding site and other features.

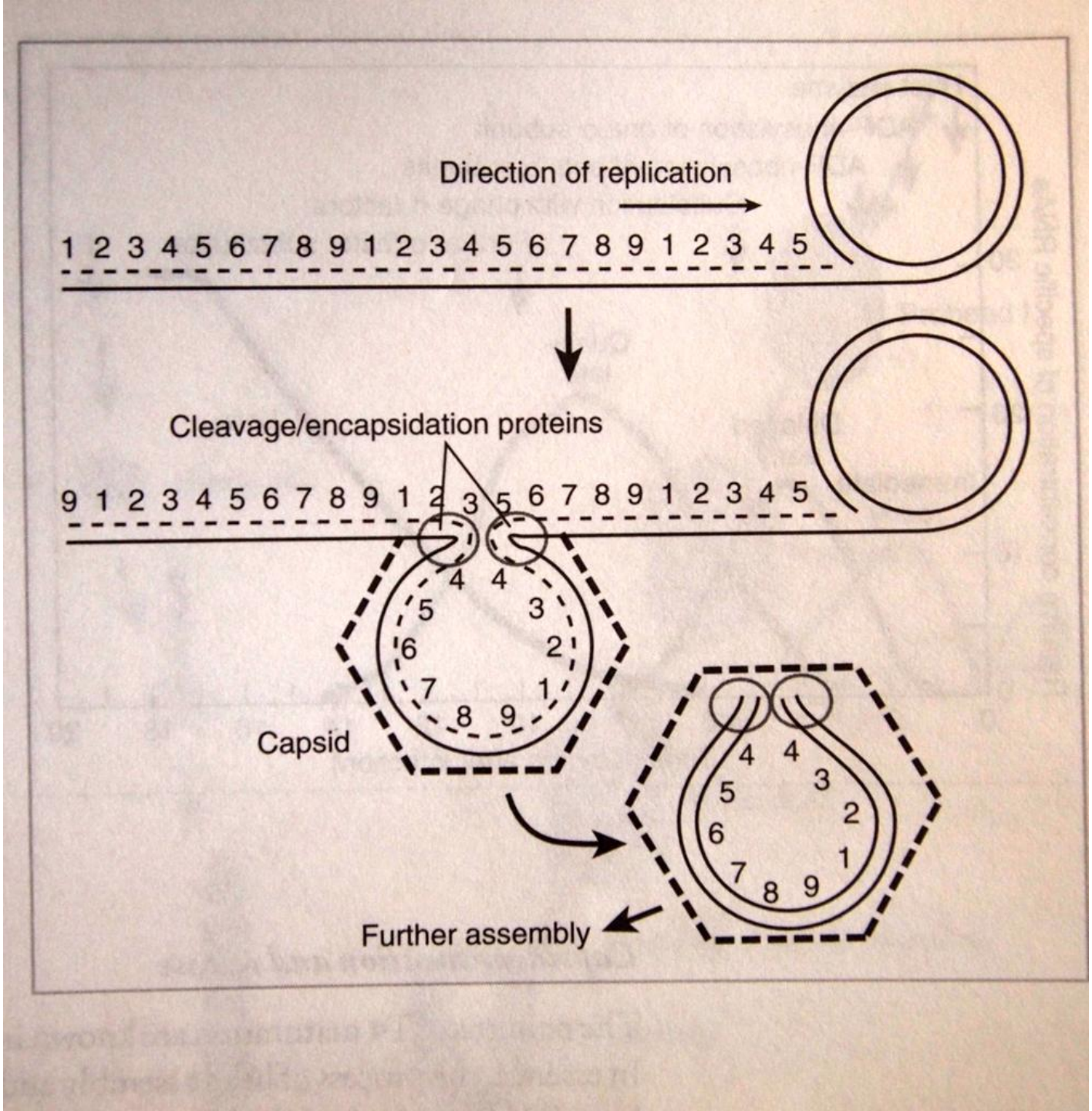
- **Terminal redundancy**

- A linear DNA molecule with the same sequence (genetic information) at each end.
- If genetic information is represented by
- ABCDEFGH
- then a terminally redundant sequence can be for instance
- ABCDEFGHAB.
- Terminal redundancy is seen in some phages (eg T2, T4) and is generated because a phage head is capable of containing a DNA molecule larger than the complete genome and packaging of DNA into phage heads is determined by the headfull. These phages also show **circular permutation**.

In T4 about 500 base pair are terminally redundant.

T4 phage is linear and show **circular permutation- starting pt. in the ln genome differs for various members of a particular viral population**

- Results from rolling cirle replication
- Packaging is seq. independent and occur by head full mechanism
- Terminal redundancy also seen in HERPES Virus


Direction of replication
1 2 3 4 5 6 7 8 9 1 2 3 4 5 6 7 8 9 1 2 3 4 5
Cleavage/encapsidation proteins
9 1 2 3 4 5 6 7 8 9 1 2 3 5 6 7 8 9 1 2 3 4 5
Capsid
Further assembly

Terminal cohesive ends

The bacteriophage λ linear genome circularizes upon injection into the cell via cohesive ends,

-12 nucleotide complimentary extensions on the 5' ends of the molecule.

early replication- theta mode-replicate bidirectionally to generate upto 20 circular progeny copies per cell.

late replication is by rolling mechanism, which generates many hundreds of progeny genomes.

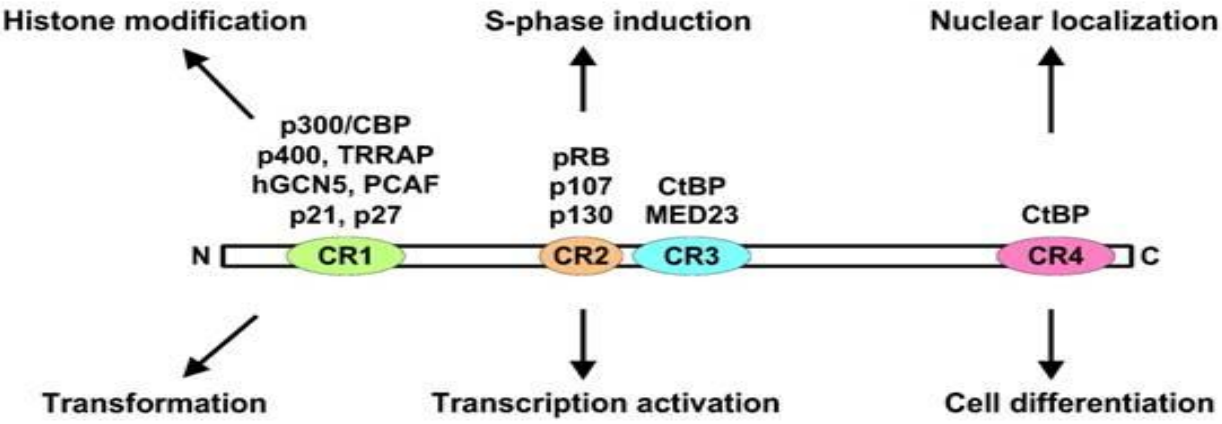
Rolling circle synthesis generate a concatemeric genome

cleaved during packaging into molecules of the correct length

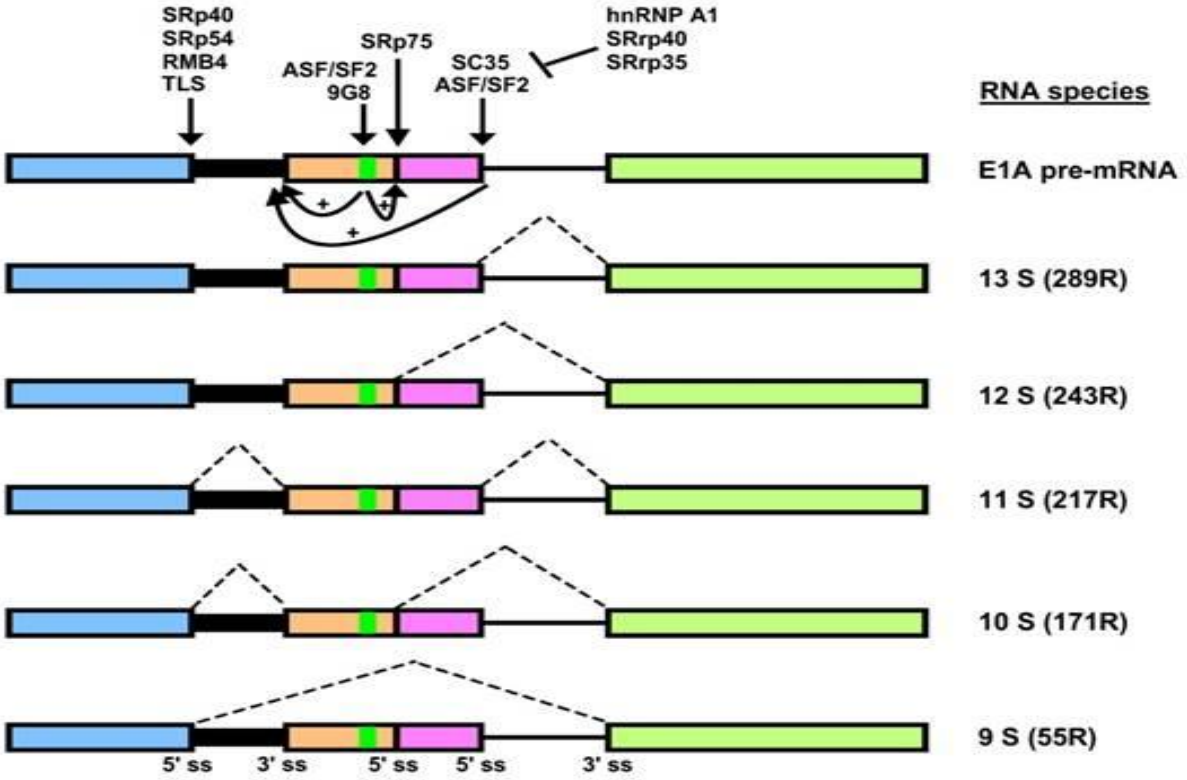
possessing the same cohesive ends as the linear genome which initiated infection.

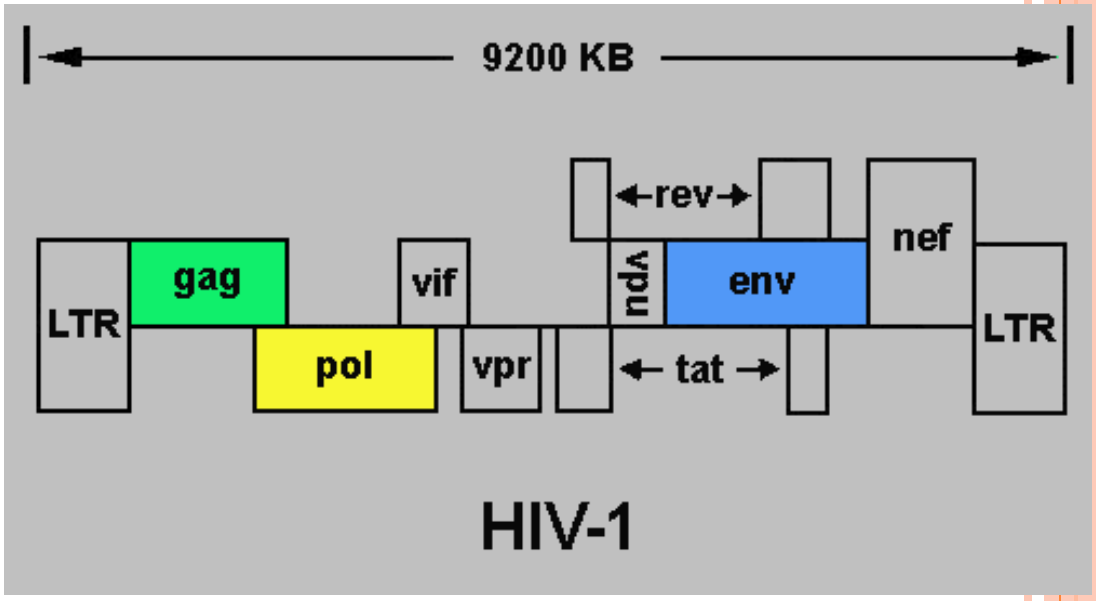
- Alternate splicing
- **Alternative splicing** (or **differential splicing**) is a process by which the coding regions of the RNA produced by transcription of a gene (a primary gene transcript or pre-mRNA) are reconnected in multiple ways during RNA splicing. The resulting different mRNAs may be translated into different protein, thus, a single gene may code for multiple proteins.
- Alternative splicing occurs as a normal phenomenon in eukaryotes, where it greatly increases the diversity of proteins that can be encoded by the genome; in humans, ~95% of multiexonic genes are alternatively spliced.
- THE MOST common METHOD is exon skipping. In this mode, a particular exon may be included in mRNAs under some conditions or in particular tissues, and omitted from the mRNA in others.

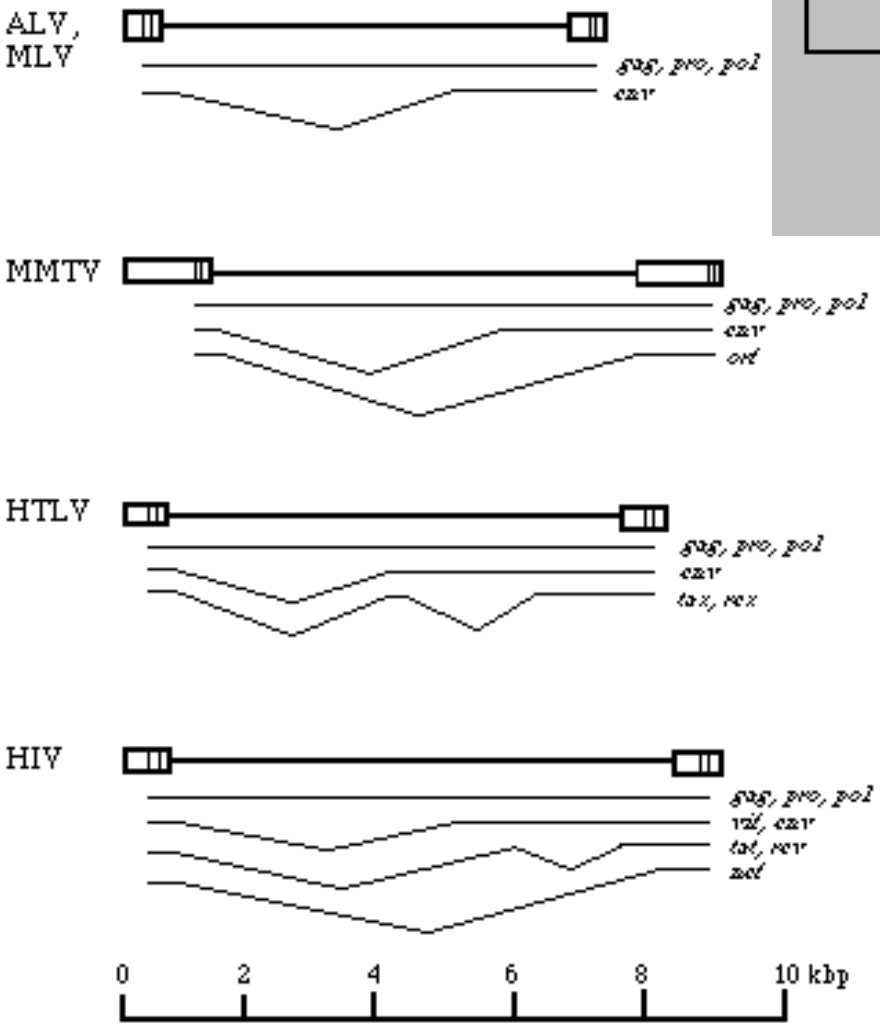
A. Adenovirus E1A-289R

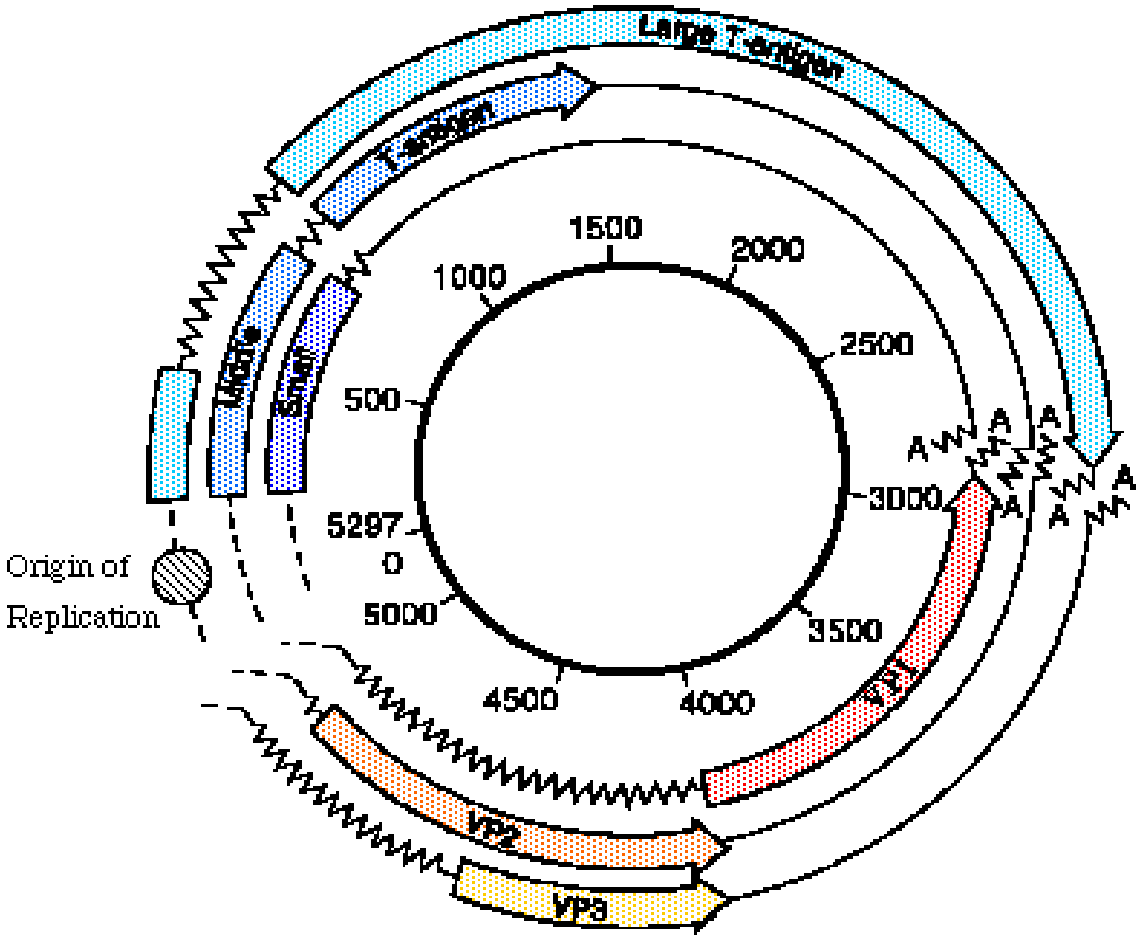


B. Adenovirus E1A RNA species




9200 KB
rev
nef
LTR
gag
vif
vpu
env
LTR
pol
vpr
tat
HIV-1


ALV, MLV
gag, pro, pol
env
MMTV
gag, pro, pol
env
ord
HTLV
gag, pro, pol
env
tax, rex
HIV
gag, pro, pol
vif, env
tat, rev
nef
0
2
4
6
8
10 kbp


Large T-antigen
T-antigen
Leader
Origin of
Replication
1500
2000
1000
2500
500
3000
5297
0
5000
3500
4500
4000
VP1
VP2
VP3
A

AMBISENSE GENOME

A SINGLE STRANDED GENOME THAT CONTAINS BOTH POSITIVE SENSE AND NEGATIVE SENSE.

ARENAVIRUS

- *Virus contains a beaded nucleocapsid with 2 single stranded RNA segments.*
- *Strands of RNA are considered –ve sense but encode genes in both directions and are thus Ambisense.*
- *The word Ambisense is derived from the Latin ambi meaning "on both sides"*
- *Half the genome is –ve polarity. But the other half is of +ve polarity and is transcribed twice.*
- *In other words, there are opening reading frames(ORFs) in both directions.*
- *First a complete transcript of the genome is made than the m RNA is transcribed from this transcript.*

PARTIALLY DOUBLE STRANDED GENOMES

COMPLETE NEGATIVE STRAND AND INCOMPLETE POSITIVE STRAND

HEPATITIS B

- *Genome is partially dsDNA that forms a covalently closed circle with 5' end of the full length minus strand which is linked to viral DNA polymerase.*
- *The genome sequence has terminii with cohesive ends that matched the uniquely located 5' ends of the two strands which overlaps by ~240 nucleotide.*
- *Complementary to the viral mRNA, the –ve strand or noncoding strand is full length; the viral +ve sense strand is shorter than full length.*

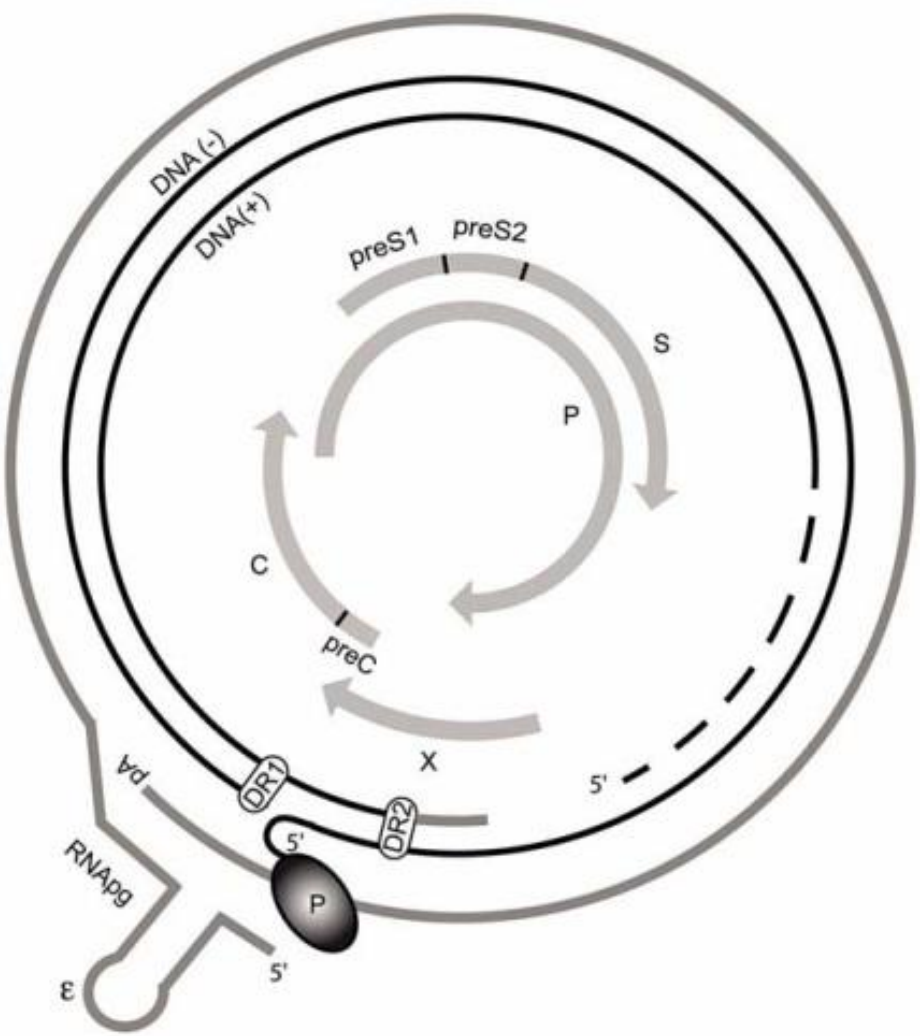
Why the (+) DNA in the virion is incomplete?

- *During (+)DNA synthesis a nucleocapsid can either migrate to the nucleus to increase the pool of DNA*

OR

- *It can undergo a maturation event that enables it to bud through a membrane containing virus envelope proteins.*
- *DNA synthesis ceases on budding, as the nucleocapsid is cut off from the pool of nucleotides in the cytoplasm.*

PARTIALLY DOUBLE STRANDED GENOMES



LONG TERMINAL **REPEATS**

RETROVIRUS

- *They are found in retroviral DNA*
- *Many linear virus genomes have repeat sequences at the ends(termini), in which case the sequences are known as terminal repeats.*
- *example:*

  LTR—PBS—PSI—gag—pol—env—LTR.

- *If the repeats are in the same orientation they are known as direct terminal repeats (DTRs).*
- *If they are in the opposite orientation they are known as inverted terminal repeats (ITRs).*

ITRs in single stranded nucleic acids are not repeated until the second strand is synthesized during replication.

- *partially transcribed into an RNA intermediate, followed by reverse transcription into RNA and ultimately dsDNA with full LTRs.*
- *The LTRs then mediate integration of the retroviral DNA via LTR specific integrase into another region of the host chromosome*
- *HIV use this basic mechanism.*
- Strong promoters are present within LTRs
- Do not encode proteins

SEGMENTED GENOME

THE GENOME IS COMPOSED OF SEPARATE SEGMENTS.

INFLUENZA VIRUS

- *Consists of ss(-)sense RNA in 8 segments*
- *The genome of each segments often codes for only one protein.*
- *They are usually found in one capsid in animal viruses and separate capsids in nonenveloped plant viruses.*
- *Mostly ss RNA genomes, rarly ss DNA eg some members of Geminiviridae contain bipartite (two cirDNA molecules)*

NON SEGMENTED GENOME

PICONARVIRUS

- *Consists of linear positive sense ssRNA of 7-8kb size.*
- *Genome is unsegmented.*
- *Has 5' VPg of RNA*
- *There is an Untranslated Region (UTR)at both ends of piconarvirus genome.*
- *The 5'UTR is longer compared to that of the 3'UTR.*
- *The 5'UTR is important in translation and the 3'UTR in negative strand synthesis.*
- *The UTR regions contains information for regulation of translation and mRNA stability.*

CAPPING AND TAILING

TOBACCO MOSAIC VIRUS (TMV)

- *TMV is a positive sense ssRNA virus that infects plants(tobacco).*
- *Exhibit capping and tailing of genome.*

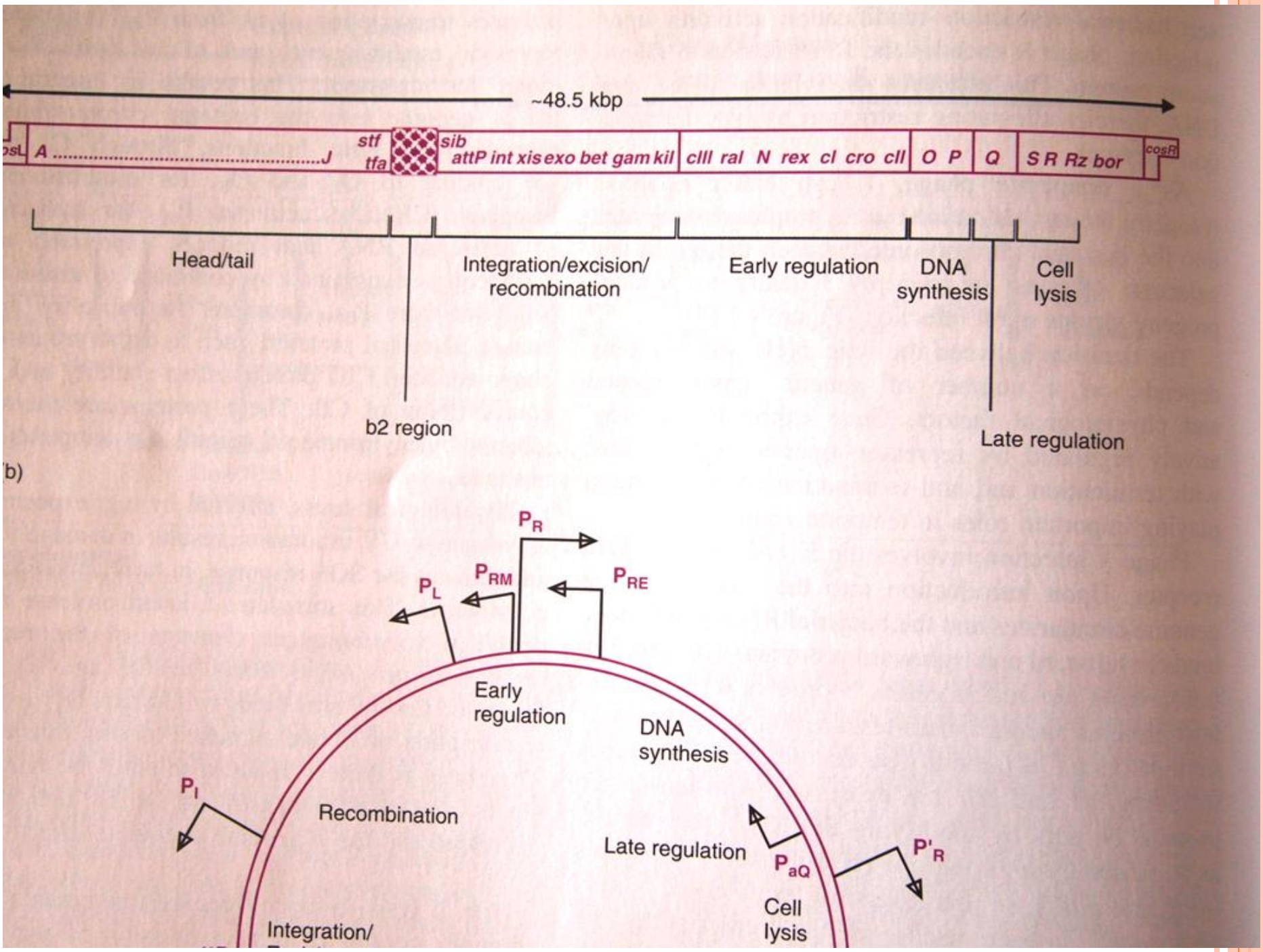
CAPPING

- *The capping enzymes adds a GTP molecule generating a phosphate link-TO phosphate. The linkage is 5' to 5' rather than the normal 5' to 3'.*
- *Next a methyl group is added by methyl transferase to nitrogen no 7 in the guanine base.*
- *The function of this cap structure seems to be to protect the mRNA from degradation. It also seems to be necessary for binding the mRNA to the ribosome during translation initiation.*

TAILING

- *Adenyl AMP residues are added to the new 3' end of the pre mRNA by an enzyme called poly A polymerase.*
- *This generates a poly A tail at the 3' end of the pre mRNA*
- *Poly A tails are present at the 3' end of all mRNAs except the mRNAs for histones.*
- *Function is unknown.*
- *The poly A tail may have something to do with protecting the 3'end of the mRNA from degradation.*

A cap and a poly A tail on a genome RNA may indicate that the molecule is ready to function as mRNA, but neither structure is essential for translation


~48.5 kbp
cosL
A ..... J
stf
tfa
sib
attP int xis exo bet gam kil
cIII ral N rex cI cro cII
O P
Q
S R Rz bor
cosR
Head/tail
Integration/excision/
recombination
Early regulation
DNA
synthesis
Cell
lysis
b2 region
Late regulation
b)
P_R
P_{RM}
P_{RE}
P_L
Early
regulation
DNA
synthesis
P_I
Recombination
Late regulation
P_{aQ}
P'_R
Cell
lysis
Integration/