Reverse transcription and integration

Lecture 9
Biology 3310/4310
Virology
Spring 2018

“One can’t believe impossible things,” said Alice.
“I dare say you haven’t had much practice,” said the Queen. “Why, sometimes I’ve believed as many as six impossible things before breakfast.”
--LEWIS CARROLL, Alice in Wonderland
Tumor virus history

- 1908 - Discovery of chicken leukemia virus, Bang & Ellerman
- 1911 - Discovery of Rous sarcoma virus, Peyton Rous (Nobel Prize 55 years later)
- Called tumor viruses
- Found to have RNA genomes
Temin’s insight

- Retroviruses caused permanent changes in cells (transformation)
- Retroviral DNA was integrated into host genome
- Became permanent part of host DNA
- Provirus
Baltimore and Temin independently discovered RT in RNA tumor virus particles

Listen to TWiV #100 (Baltimore) for more insight
Reverse transcriptase

- Enzyme that countered *Central Dogma*:
  
  DNA $\Rightarrow$ RNA $\Rightarrow$ protein

- Retroviruses got their name because of their ability to reverse the flow of genetic information

- RT discovery revolutionized molecular biology
Viruses with RT

- **Retroviruses**
  - Poliovirus
  - Influenza virus
  - VSV
  - Hepatitis B virus
  - Reovirus

**Steps:**

1. + DNA → ± DNA
2. ± DNA
3. ± RNA
4. + mRNA → ± RNA
5. ± RNA → − RNA
6. + RNA → − DNA
Simple genome (ALV)

Proviral DNA

Core
Enzymes
Envelope

LTR 3'

gag pol env LTR 5'

MA CA PR RT IN SU TM

5'

5'

Genome expression

Genomic RNA, Gag-Pol mRNA, pre-mRNA

5'

gag  
 pol  
 env  

A_{n}A_{OH}^{3'}

Translation

Mature, viral PR-processed proteins

Gag-Pol precursor

RT α

β

Mature, viral PR-processed proteins

Singly spliced Env mRNA

Envelope precursor

SU TMI

A_{n}A_{OH}^{3'}
Reverse transcriptase

- Primer can be DNA or RNA
- Template can be RNA or DNA
- Only dNTPs, not rNTPs, are incorporated
RT

- Bacteria and Archaea have RT activity
- Therefore RT evolved before the separation of Archaea, bacteria, and eukaryotes
- RT might be the bridge between early RNA world and modern DNA world
- RT also in HBV, Caulimoviridae
Sequence relationships among polymerases

- Gly-Asp-Asp in (+) strand RNA polymerases
- Asp-Asp in RT, segmented (-) strand polymerases
- Gly-Asp-Asn in nonsegmented (-) strand polymerases
RNAse H: A second activity of RT

- Cleaves RNA only when in duplex form
- RNA can be in RNA:RNA or RNA:DNA duplexes
- Makes endonucleolytic cleavages
- Produces short oligonucleotides with 5’-phosphate, 3’-OH
DNA synthesis is slow (4 h per 9 kb genome) and error prone (1 misincorporation per $10^4$ to $10^6$ nt)
Reverse transcriptase has revolutionized molecular biology. Which statement about the enzyme is not correct?

A. RT is unique to retroviruses
B. RT is packaged in the retrovirus particle
C. The RT protein also has RNAse H activity
D. The name of the enzyme comes from its ability to reverse the flow of genetic information
E. Might have bridged the ancient RNA world and the DNA world
Coated with NC protein

50-100 molecules RT per virus particle
RNA dimer

- Explains why retroviruses are relatively resistant to UV and ionizing radiation
- Two copies of all genes
- Copy-choice rebuilds one functional genome
DNA synthesis: cytoplasmic

Initiation of (−) strand DNA synthesis

1. Viral injection into the cytoplasm
2. Uncoating of the viral genome
3. Expression of viral enzymes

Virology Lectures 2018 • Prof. Vincent Racaniello • Columbia University
First template exchange

ppt = polypurine tract
(+) strand DNA synthesis

PPT  U3  R  U5  Modified base

PPT  U3  R  U5

PPT  U3  R  U5  PBS

(+) strong-stop DNA
Second template exchange is facilitated by annealing of PBS sequences
Which of the following steps occur during reverse transcription of retroviral genomic RNA?

A. Priming of (-) DNA synthesis by tRNA  
B. Two template exchanges  
C. Degradation of the viral RNA by RNAse H  
D. Generation of two LTRs  
E. All of the above
Preference for integration into DNA sequences that are wrapped around a nucleosome
• One DNA produced from two RNAs by RT
• Strong promoter (the LTR) built during RT
• Proviral DNA directs the host transcription machinery to synthesize many copies of viral mRNA
• Viral mRNA is translated into viral proteins OR encapsidated into virus particles

There is no viral DNA replication and no viral RNA replication
• No mechanism for precise excision of integrated provirus
• Only way out of genome is transcription by host RNA pol II
• Genomes are littered with ancient and modern retroelements
Retroelements

- Sequences that move in the genome via RT
- Proviral DNA integrated into the germline = endogenous retroviruses, ERV
- Often replication-defective
- ~42% of human genome comprises mobile genetic elements, including endogenous proviruses and other retroelements
Retroelements in the human genome

Endogenous retrovirus

Retrotransposons *

LINEs

SINEs

Processed pseudogenes

* Likely progenitors of retroviruses
Rescue of an endogenous human retrovirus

- HERV-K, infected human ancestors <1 Myr ago
- Repaired mutations - known HERVs are not infectious
Endogenization

[Diagram showing genomic integration]

Virology Lectures 2018 • Prof. Vincent Racaniello • Columbia University

Syncytins: Exapted retroviral env

A

syncytin-1 (HERV-W)
syncytin-2 (HERV-FRD)

Human
Chimpanzee
Gorilla
OrangUtang
Gibbon
Old World Monkeys
New World Monkeys
Prosimian

60 40 30 20 10 0
Million years

B

fetal villi
uterus
fetal vessel
maternal blood
cytotrophoblast (CT)
syncytiotrophoblast (ST)
extravillous trophoblasts (EVT)
uterine artery
A retrovirus makes chicken eggshells blue

Which of the following statements about retroelements is not correct?

A. There are many copies in eukaryotic genomes  
B. They are currently entering the Koala germline  
C. Those in the human genome produce infectious viruses  
D. They can be beneficial  
E. None of the above
Hepadnaviridae

Viral particle
Viral DNA
Polymerase (P)
Capsid
Large (L)
Medium (M)
Small (S)
42 nm

Incomplete particles

20 × 20–200 nm

15–25 nm
No genome integration
1. **Initiation**

   - dr2
   - dr1
   - AₙA$_{OH}$³'

2. **First template exchange**

3. **RNA template removed as DNA synthesis proceeds**

4. **RT**

   - RNA primer
   - ATGTGGGGAGAG

   - DR1
   - DR2
   - 3' CAGAAA UACACCCCUUC
   - 5' CAGAAA UACACCCCUUC

   - TP