Evolution

Lecture 21
Biology W3310/4310
Virology
Spring 2014

Anything produced by evolution is bound to be a bit of a mess.
SYDNEY BRENNER

Around here, it takes all the running you can do just to stay in the same place.
LEWIS CARROLL
Alice in Wonderland
Viral evolution: The constant change of a viral population in the face of selection pressures

- Viral populations display spectacular diversity - why they are successful

- Sources of diversity:
  - Mutation
  - Recombination

- If the population can’t change or adapt, it disappears
Modern virology has provided a window on the mechanisms of evolution

- As host populations grow and adapt, virus populations are selected that can infect them
  
  - *New viral populations emerge every day*

- It also works the other way
  
  - *Viral populations can be significant selective forces in the evolution of host populations*

- If a host population cannot adapt to a lethal virus infection, the population may be exterminated
The public is constantly confronted with the reality of viral evolution (even if they don’t believe in evolution)

- New viral diseases: AIDS, West Nile virus in the US, HCV, ebolavirus
- Regular bouts every year with influenza and common cold viruses
- Drug resistant HIV

Simple fact: viruses evolve faster than many can comprehend
Viral evolution

- Defined in terms of a population, not an individual virus particle
- Viral populations comprise diverse arrays of mutants that are produced in prodigious quantities
- It is misleading to think of an individual particle as representing an ‘average’ virion for that population - virologists are population biologists whether they know it or not
Four main drivers of virus evolution

• Large numbers of progeny
• Large numbers of mutants
• Quasi-species effects
• Selection
Virus-infected cells produce large numbers of progeny

The interface of host defense and virus replication is fertile ground for selection and evolution

**Table 10.1  *In vivo* dynamics of human immunodeficiency virus and hepatitis B virus**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value in:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hepatitis B virus</td>
</tr>
<tr>
<td>Virus in plasma</td>
<td></td>
</tr>
<tr>
<td>Half-life</td>
<td>24 h</td>
</tr>
<tr>
<td>Daily turnover</td>
<td>50%</td>
</tr>
<tr>
<td>Total production in blood</td>
<td>&gt;10^{11}</td>
</tr>
<tr>
<td>Virus in infected cell</td>
<td></td>
</tr>
<tr>
<td>Half-life</td>
<td>10–100 days</td>
</tr>
<tr>
<td>Daily turnover</td>
<td>1–7%</td>
</tr>
</tbody>
</table>
Replicating viruses produce large numbers of mutant genomes

- Evolution is possible only when mutations occur in a population
- Mutations are produced during copying of any nucleic acid molecule
RNA viruses

• The masters of error-prone replication: RNA polymerases cannot correct errors

• Average error frequency: 1 in $10^4$ or $10^5$ nucleotides polymerized

• In a 10 kb RNA virus genome, a mutation frequency of 1 in $10^4$ results in about 1 mutation per genome
DNA viruses

- Genome replication not as error prone as RNA viruses
- Proofreading
- Most DNA viruses generate less diversity, evolve slower than RNA viruses
- Different lifestyle from most RNA viruses
  - Narrow host range
  - Persistent infections common
The quasispecies concept

• Analysis of an RNA bacteriophage population (Qβ):

“A Qβ phage population is in a dynamic equilibrium with viral mutants arising at a high rate on the one hand, and being strongly selected against on the other. The genome of Qβ cannot be described as a defined unique structure, but rather as a weighted average of a large number of different individual sequences.” E. Domingo, D. Sabo, T. Taniguchi, C. Weissmann. 1978. Nucleotide sequence heterogeneity of an RNA phage population. Cell 13:735-744.

• This discovery was far ahead of its time, not appreciated by most virologists

• Virus populations exist as dynamic distributions of nonidentical but related replicons, called quasispecies
Viral quasispecies

this

not this
Viral infections are initiated a population of particles, not a single virion

The large number of progeny produced are complex products of intense selective forces inside the host

The survivors that can re-infect a new host reflect the intense selective forces outside the host
The myth of consensus genome sequences

• For a given RNA virus population, the genome sequences cluster around a consensus or average sequence, but virtually every genome can be different from every other.

• It is unlikely that a genome with the consensus sequence is actually replicating in the population.
Selection

- Selection favors both the creation of diversity and single dominant mutations
  - Counterintuitive - the drive for survival must result in diverse populations as well as selected mutations
Selection

- *Survival of the fittest*: A rare genome with a particular mutation may survive a selection event, and this mutation will be found in all progeny genomes.

- *Survival of the survivors*: However, the linked, but unselected mutations, get a free ride.

- Consequently, the product of selection after replication is a new, diverse population that shares only the selected mutations.
Selection for survival inside a host

- At the end stage of AIDS, billions of virions present engage Cxcr4
- Virions passed to a new host engage Ccr5
- Progeny genomes have passed through a bottleneck
- Virions that ultimately devastate the immune system are not the most fit for infection of new hosts
Diversity is selected

- Mutations in viral polymerases that reduce the frequency of incorporation errors
  - Do not have a selective advantage when wild type and anti-mutators are propagated together
  - Lower rates are neither advantageous nor selected in nature
  - Mutants are often less pathogenic
- High mutation rates are selected during virus evolution: mutation is good for viral populations

http://www.virology.ws/2009/05/15/increased-fidelity-reduces-viral-fitness/
Error threshold

- Mutation is a powerful advantage, but selection and survival balances genetic fidelity and mutation rate

- This limit is called the **error threshold**
  - Exceed it: loss of infectivity
  - Below it: cannot produce enough mutations to survive selection

- RNA viruses: evolve close to their error threshold

- DNA viruses: evolve far below their error threshold
Error threshold

• Expose a cell culture infected with a DNA virus to a base analog such as 5-azacytidine

• 5-azacytidine is incorporated as a C, but templates as a T (G to A transitions)

• Mutation rate among viral progeny increases several orders of magnitude

• When a similar experiment is done with an RNA virus, the error frequency per genome increases only two- to threelfold at best - cannot make any more mutations
Error threshold

antiviral ribavirin and poliovirus
Genetic bottlenecks

- Extreme selective pressures on small populations that result in loss of diversity, accumulation of non-selected mutations, or both

- A single RNA virus plaque is picked and expanded

- Next, a single plaque is picked from the expanded stock

- The process is repeated over and over
Gene&c
  
  Gene&c

Genetic bottlenecks

• After about 20-30 cycles of single-plaque amplification, many virus populations are barely able to grow

• They are markedly less fit than the original population

• The environment is constant, and the only apparent selection is that imposed by the ability of the population of viruses from a single plaque to replicate

• Why does fitness plummet?
Genetic bottlenecks

- The bottleneck arises by restricting further viral replication to the progeny found in a single plaque
  - A few thousand progeny viruses derived from a single founder virus
Genetic bottlenecks

• The answer lies in a phenomenon dubbed Muller’s ratchet: Small, asexual populations decline in fitness over time if the mutation rate is high.

• Replicating RNA viruses produce many mutations; they survive close to their error threshold.

• By restricting population growth to serial single founders (the bottleneck) under otherwise nonselective conditions, so many mutations accumulate (exceed the threshold) that fitness decreases.
The ratchet metaphor: each of the new mutations works like a ratchet, allowing the gear to move forward, but not backward

Each round of error-prone replication works like a ratchet, “clicking” relentlessly as mutations accumulate at every replication cycle
Mutations that improve replication

Increased replicative fitness

Large population passages

Consensus sequences

Bottlenecks (small population passages)

Repeated bottlenecks

Accumulation of mutations

Decreased replicative fitness
Fitness decline compared to initial virus clone after passage through a bottleneck

<table>
<thead>
<tr>
<th>Virus</th>
<th>No. of bottleneck passages</th>
<th>% Decrease in fitness (avg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$6 (bacteriophage)</td>
<td>40</td>
<td>22</td>
</tr>
<tr>
<td>Vesicular stomatitis virus</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>Foot-and-mouth disease virus</td>
<td>30</td>
<td>60</td>
</tr>
<tr>
<td>Human immunodeficiency virus</td>
<td>15</td>
<td>94</td>
</tr>
<tr>
<td>MS2 (bacteriophage)</td>
<td>20</td>
<td>17</td>
</tr>
</tbody>
</table>

Bottlenecks in the real world?

- Infection by a limited virus population and subsequent amplification are often found in nature
  - *Small droplets of suspended virus during aerosol transmission*
  - *Activation of a latent virus from a limited population of cells*
  - *Small volume of inoculum introduced in infection by insect bites*

- How do infections that spread by these routes escape Muller’s ratchet?
Avoiding the ‘ratchet’

- Subject a more diverse viral population to serial passage
  - Don’t pick a single plaque, pool several plaques

- More diversity in the replicating population facilitates construction of a mutation-free genome by recombination or reassortment, removing or compensating for mutations that affect growth adversely
By exchange of genetic information

Sick virus 1

Sick virus 2

Healthy viral recombinant
Avoiding the ‘ratchet’

• Even if such a recombinant is a rare species, it has a powerful selective advantage for growth
  - Its progeny will ultimately take over the population

• The message is simple: Diversity of a viral population is important for the survival of individual members
  - Remove diversity, and the population suffers
Exchange of genetic information

• Allows the construction of viable genomes from debilitated ones and avoids Muller’s ratchet
• A similar mechanism is reassortment among genomic segments when cells are co-infected with segmented RNA viruses
• It is an important source of variation, as exemplified by orthomyxoviruses and reoviruses
Selection: Genetic shift & drift

- Selection of viral mutants resistant to elimination by antibodies or cytotoxic T cells inevitable when sufficient virus replication occurs in an immunocompetent individual.

- Drift - diversity arising from copying errors and immune selection - may occur each time a genome replicates.

- Shift - diversity arising after recombination or reassortment - is relatively rare.
Influenza A viruses are classified by antigenic composition, by serologic testing of HA and NA.

- Combinations of H and N are called HxNy.
- $x = 1-18; \ y = 1-11$
- H1-17 can infect birds; H1, H2, H3 can infect and transmit between humans.
Antigenic drift: Influenza virus
Exchange of genetic information
Selection: Is virulence a positive or negative trait?

- Idea: increased virulence reduces transmissibility because hosts die faster, reducing exposure to uninfected hosts.
- Expectation: all viruses evolve to be maximally infectious and avirulent.
- But this is not observed.
- Persistent infections lie dormant for years, then kill host at end stage.
- Virulent viruses for one species may be maintained as asymptomatic infections in another.
- For some diseases, increased virulence increases transmissibility and is selected for in natural infections.
An experiment in virus evolution

• In 1859, the European rabbit was introduced to Australia for hunting purposes

• Lacking natural predators, it reproduced to plague proportions in a short time
An experiment in virus evolution

• The myxoma leporipox virus was released in Australia in the 1950s in an attempt to rid the continent of the rabbits

• The natural host of myxoma virus is the cottontail rabbit, the brush rabbit of California, and the tropical forest rabbit of Central and South America

• The virus is spread by mosquitoes; infected rabbits develop superficial warts on their ears

• European rabbits are a different species, infection is 90-99% fatal
An experiment in virus evolution

- In the first year, the released virus was efficient in killing rabbits with a 99.8% mortality rate
- By the second year the mortality dropped to 25%
- In subsequent years, the rate of killing was lower than the reproductive rate of the rabbits, and hope for 100% eradication was dashed
An experiment in virus evolution

- Both rabbits and viruses produce large numbers of offspring
- The virus evolved to kill fewer rabbits and to extend the life of lethally infected rabbits so that the virus could overwinter and spread in spring mosquitoes
- The rabbits evolved to become more resistant or tolerant of the virus
- As predicted for an evolving host coming to an equilibrium with the pathogen
Why do viruses cause disease?

- Virulence may be ‘accidental’
  - *High levels of replication needed for transmission (but compare poliovirus vs norovirus)*
  - *Immune modulation*

- Bacteria: virulence as collateral damage (virology.ws/2014/02/07/why-do-viruses-cause-disease/)

- Host changes and virulence
The origin of viruses

• Virus-like elements preceded cellular life
• Viruses arose from cells by reductive evolution
• Viruses arose from cells and gained genes (‘pickpockets’)
• There is no fossil record, and few viral stocks more than 80 years old*
• Bioinformatics has provided insight
Very large viral DNA genomes

- *Mimivirus* (1.2 Mbp), *Pandoravirus* (2.5 Mbp) genomes are largely dark matter

- Arose from fourth domain of life?

http://www.twiv.tv/2013/12/01/twiv-261-giants-among-viruses/
Origins of DNA viruses

- By relating timescale of herpesviral genome evolution with that of hosts, believe that three major groups of herpesviruses (alpha, beta, gamma) arose \(~180-220\) million years ago.
Endogenous viruses - retrovirus and otherwise

- ERVs - date by DNA mutation rate in genome e.g. HERV-H
- Phylogenomics
- Unexpected endogenous viruses
Human viruses

- All known types of viruses likely evolved long before humans appeared on Earth
- All human viruses have therefore evolved from animal viruses
Origins of smallpox virus

- Sequence analysis of 45 isolates: genomes can be organized into 3 clades, which cluster geographically (West Africa, South America, Asia)
- Gene content is constant; lack of diversity indicates recent introduction into humans
- Only member of poxvirus family with credible homology to smallpox is a gerbil poxvirus
- Perhaps human smallpox virus arose after a zoonotic infection from infected gerbils
Origins of measles virus

• Measles virus is closely related to rinderpest virus, a bovine pathogen

• Probably evolved from an ancestral rinderpest virus when humans first began to domesticate cattle

• Established in the Middle East ~5,000 years ago, when human populations began to congregate in cities

• Spread around the world by colonization and migration, reaching Americas in 16th century and destroying native Americans
Evolution of new viruses

• Assumption: new viruses can only arise from viruses that are now in existence, not *de novo*

• What is the number of all possible mutations of a viral genome?

• Sequence comparisons of several RNA virus genomes have demonstrated that well over half of all nucleotides can accommodate mutations
Evolution of new viruses

- For a 10 kb viral genome, $4^{5000}$ sequences
- Deletions, recombination, and reassortment increases the numbers
- $\sim 4^{135}$ atoms in the visible universe
The fundamental properties of viruses constrain and drive evolution

- Despite many rounds of replication, mutation, selection, we can recognize a herpesvirus or influenza virus genome by sequence analysis.
- Viral populations often maintain master or consensus sequences, despite opportunities for extreme variation.
- How is stability maintained?
- All viruses share fundamental characteristics that define and constrain them.
- Viruses that can function within the constraints survive.
Constraining viral evolution

- Extreme alterations in viral consensus genome do not survive selection
- The viral genome is one constraint
  - DNA cannot become RNA, or vice versa
  - Replication strategy - cannot change; consider interaction with host proteins
- Physical nature of capsid
  - Icosahedral capsids: defined internal space, fixes genome size
- Selection during host infection
  - A mutant too efficient in bypassing host defenses will kill host, suffer the same fate as one that does not replicate efficiently enough
Food for thought

Imagine what a virus can do with 8 million years

5 days

8 million years