Key concepts in this lecture

• Quasispecies
• Viral Fitness
• Fitness Landscapes
• Using fitness landscapes to think about how viral populations evolve

How is viral genetics different?

• Genomes are haploid (except retroviruses)
• Population size is enormous compared to other organisms (can be $10^6$ to $10^{10}$ infectious units/ml), and generation time is fast
• Frequent population bottlenecks
• Much higher mutation rate (RNA viruses)

The mutation rate for RNA viruses is high because RNA-dependent polymerases lack proof-reading

• Mutation rate for eukaryotic cells is about $10^{-8}$ to $10^{-10}$ errors/bp-cycle
• Mutation rate for RNA viruses is about $10^{-4}$ to $10^{-5}$ errors/bp-cycle
• So, the mutation rate for RNA viruses is one thousand to one million times higher than the genomes of their host cells
Consequences of high mutation rates

The chimpanzees genome differs from the human genome by about 2% in the ~5 million years since the two species shared a last common ancestor.

Polio virus acquires 2% nucleotide divergence in the 5 days that it takes the virus to go from the mouth to the gut in one individual.

Mutation rate ($\mu$) versus mutation frequency

Mutation rate is the number of mutations per cycle of replication

How many mutations occur

Mutation frequency is the composite of the mutation rate, the number of cycles, and the fixation rate of mutations

How many mutations you observe

Mutation rate = mutation frequency in a single cycle in the absence of selection

Consequences of high mutation rates

$10^{-4}$ to $10^{-5}$ errors per base per cycle

$X \times 10^6$ bases (average genome size)

= nearly every genome contains at least one mutation relative to its immediate parent

The quasispecies
The viral quasispecies

Viral quasispecies are closely related mutant and recombinant viral genomes subjected to continuous genetic variation, competition, and selection.

The viral quasispecies

Viral quasispecies are genetic variants organized around a set of genotypes known as the master sequence

The viral quasispecies

(Another image as a way to think about it)

Viral quasispecies are genetic variants organized around a set of genotypes known as the master sequence.
Diversity increases with time
Phylogenetic Diversity of HIV-1

0.0024 substitutions per bp per year

Korber et al. Science 2000 Jun 9;288(5472):1789-96

Global Distribution of HIV-1 Clades

How would you interpret single clades elsewhere in the world?

All clades are found in west-central Africa
How would you interpret that?

Fitness and the viral quasispecies

Fitness quantifies the adaptation of an organism to a given environment

For a virus, fitness measures its ability to produce infectious progeny

Fitness is always a relative value i.e. fitness relative to other variants
Fitness and the viral quasispecies

At steady-state (no selection) the master sequence is stable because it represents the most fit variants.

But, the amount of overall variation (diversity) will increase.

Fitness as a function of genotype can be represented as a "fitness landscape".

Fitness landscapes include peaks and valleys.

Selection drives populations up local fitness peaks.
Example of a simple fitness landscape with a local maximum

Fitness landscapes are multidimensional

The quasispecies changes because the fitness landscape has been altered

what was the peak in one condition becomes a valley in another example?
The viral quasispecies

Different selection pressures will lead to different master sequences.

Selection pushes the quasispecies up the fitness peak

but, most mutations will push it back down the peak. Why?

FREQUENCY

<table>
<thead>
<tr>
<th>V1</th>
<th>6/85</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FREQUENCY</td>
</tr>
<tr>
<td></td>
<td>FREQUENCY</td>
</tr>
<tr>
<td></td>
<td>FREQUENCY</td>
</tr>
<tr>
<td></td>
<td>FREQUENCY</td>
</tr>
<tr>
<td></td>
<td>FREQUENCY</td>
</tr>
<tr>
<td></td>
<td>FREQUENCY</td>
</tr>
<tr>
<td></td>
<td>FREQUENCY</td>
</tr>
<tr>
<td></td>
<td>FREQUENCY</td>
</tr>
<tr>
<td></td>
<td>FREQUENCY</td>
</tr>
</tbody>
</table>

V1 6/85
V2 3/86
V3 6/86
V4 12/87

Selection time
location
different master sequences
new master sequence

 Why?
The high viral error rate pushes the population down the fitness peak, but makes it more likely that new adaptive peaks will be populated.

In a constant environment, viral populations increase fitness exponentially followed by stochastic variations around a plateau.

Why?

Theoretical consequences of the viral quasispecies

- Mutational coupling and group selection
- Error thresholds and error catastrophe
- The link between genome size and mutation rate
The mutation rate of most RNA viruses is just below their error threshold.

Small genome size and errors

• Most RNA viruses have small genomes because larger genomes would be defective due to too many mutations per genome

• Selection for fast replication?

• Eigen’s paradox

Viral genetics concepts you should know

• What is a quasispecies?
• How does error rates and population size contribute to a quasispecies?
• What is a fitness landscape and how can you use it to describe a quasispecies?
• What are the consequences of high diversity? How do selection pressures affect the quasispecies?

Confused or just want to know more: try reading this review: Quasispecies Theory and the Behavior of RNA Viruses http://dx.doi.org/10.1371/journal.ppat.1001005
Paper for discussion next time: Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population

• What is the hypothesis to be tested here?
• How did they show it?
• How do the authors interpret their results in terms of quasispecies theory?
• Do you believe it?