Variation, Varitions, Varytions, Vary Notions

36-149 The Tree of Life

Christopher R. Genovese
Department of Statistics
132H Baker Hall x8-7836
http://www.stat.cmu.edu/~genovese/

.

A "Random" Activity

1. Select a partner. The two of you should pick a single number between 1 and 25. Write this number down.
2. I will give you a starting list of 10 pairs of random numbers. Record the numbers in a column, keeping the pairs distinct: for example, (10,17), (9,3), ...
3. For each generation, do the following:
   A. Choose the five pairs whose *first* number is closest to your chosen number. Here closeness is meant in a circular sense as though the 1 to 25 were placed around the face of a clock. That is, 24 is closer to 1 and to 20 than to 12.
   B. The five surviving pairs will now reproduce. Each pair will produce two "offspring" as follows. Given a pair (x, y) select two numbers from the random increments list, call these a and b. The offspring pair is $(x + a, y + b)$. However, if the sum goes above 25 or below 1 it should "wrap around" as on a clock. Do this twice for each pair.
   C. Record your new pairs in a column next to the lists from the previous generation.
   D. Repeat ad nauseum.

Variation and Evolution

Evolution proceeds in two steps

1. Variation arises within a population.
2. The proportion of variants within the population changes across generations.

Because only heritable variation is passed, genetic variation is of particular importance.

Key Idea: Mechanisms of evolution operate on the variation in the population.

Important questions:

- What are the sources of variation?
- How much variation is there? What kinds?
- How does it change as evolution proceeds?
- Is there enough variation to produce novel structures?

Sources of Phenotypic Variation

- Variation in genotype

  Many – but not all – differences in genes lead to different traits.

- Variation in environment

  A phenotype is a product of both genotype and environment.

  Even identical twins can be distinguished by those who know them.

  Examples: environmental sex determination, dung beetle horns, alternate leaf forms

- Maternal (and sometimes paternal) effects

  In most species, the female gamete (e.g., the egg) is larger than the male gamete (e.g., the sperm) and provides the environment for the developing embryo.

  Genetic and non-genetic materials in this environment guide and shape development of the offspring.

Sources of Phenotypic Variation (cont'd)

- Maternal (and sometimes paternal) effects (cont'd)

  The structure and composition of this environment can greatly influence form, scale, and traits.

  Examples: egg size and larval form in sea urchins, *Drosophila*, ...

- Epigenetic effects (*epi-* on, in addition, after)

  Biochemical processes can modify (really, accent) the nucleotides on a DNA strand without changing the DNA sequence.

  These "accents" can influence gene expression and the pattern of some epigenetic changes is *heritable*.

- Others

  Culture, life history, ...

Sources of Genetic Variation

- Mutation
- Recombination
- Selfish genetic elements (e.g., transposons)
- Gene flow
- Hybridization
- Lateral Gene Transfer
- Sex

Viruses

Viruses are sub-microscopic, (obligate) intracellular parasites.

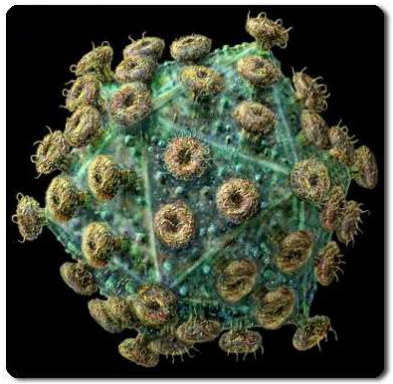
They infect virtually all forms of life, including animals, plants, fungi, and bacteria.

Parasites generally obtain resources from their host, but viruses depend totally on their host cell's machinery to replicate.

But they do replicate, and in the process, evolve in recognizable lineages.

This challenges our definition of life.

Virus Structure

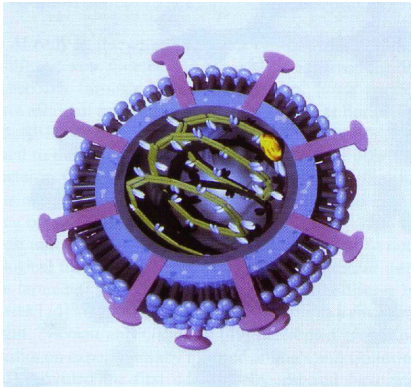


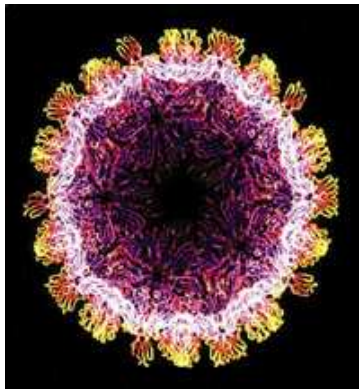
A single virus particle is called a *virion*.

Each virion consists of

- an outer protein shell, called a *capsid*,
- an inner core containing genetic material, either RNA or DNA, and some additional proteins to initiate replication.
- for some viruses, macro-molecules that protrude through the capsid and are used for binding to and penetrating cells.

Viruses are classified by the nature of their genetic material. Most have only a small number of genes.





Virus Life Cycle

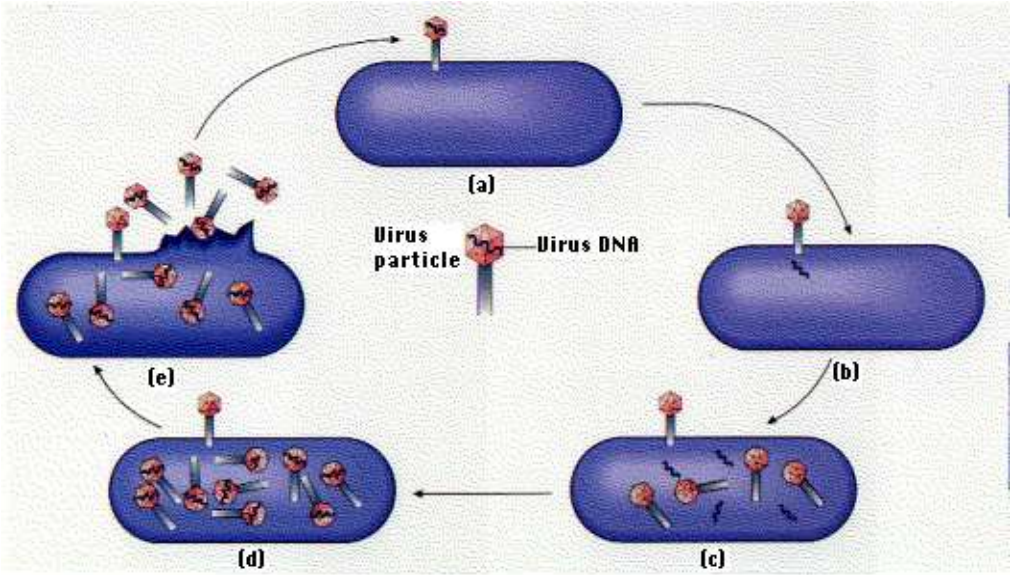
In general terms, as follows:

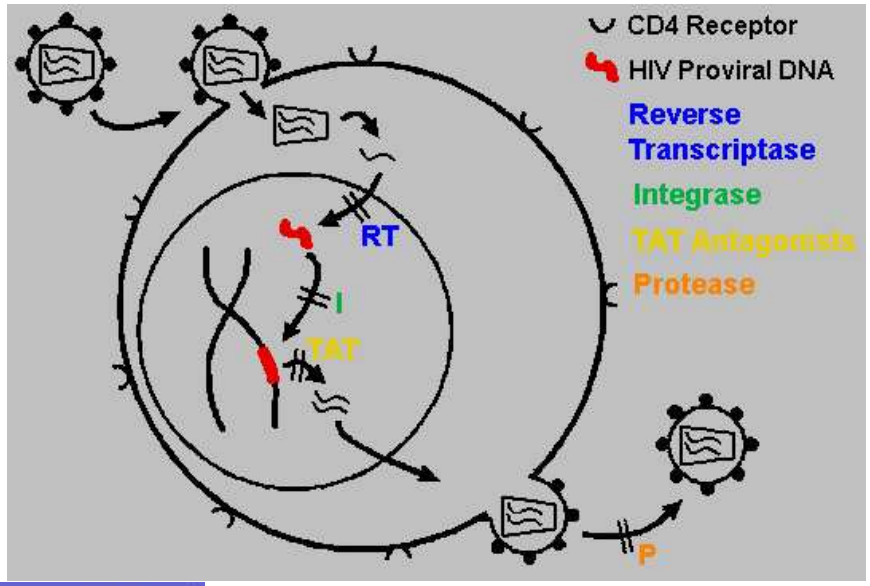
1. Virion attaches to the surface of a potential host cell, usually binding to a specific type of molecule.
2. The virion is taken into the cell.
3. Cellular enzymes remove the capsid and release the viral core contents into the cell.
4. The viral genes arrange to get themselves expressed and then synthesize
   i. new copies of the genome
   ii. the proteins needed to construct new virions.
5. The newly constructed virions escape the cell.

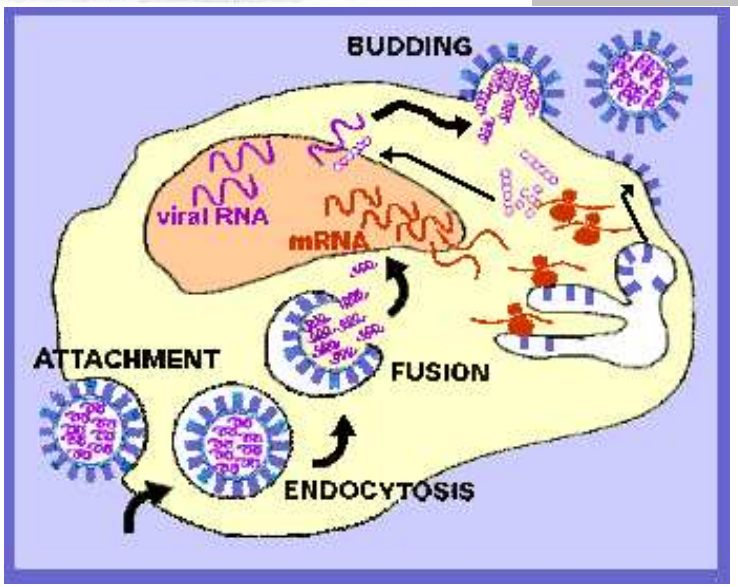
   Sometimes they burst out en masse, killing the cell outright; sometimes they emerge a few at a time; sometimes they become latent infections.

The details differ for each type of virus.

Virus Life Cycle (cont'd)





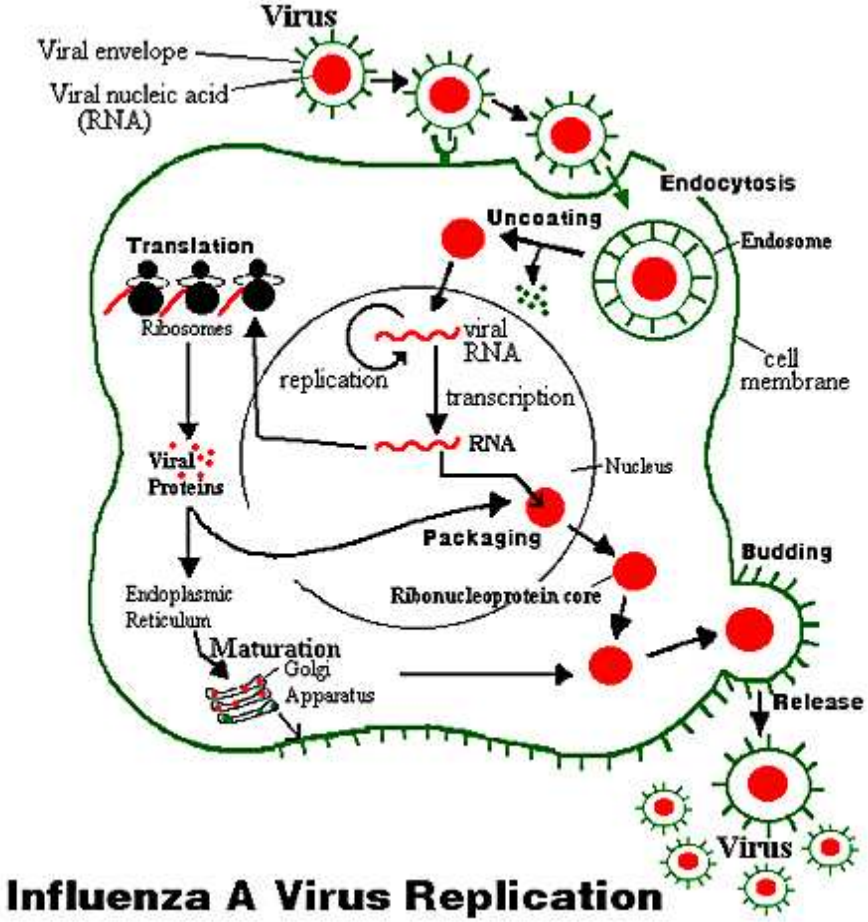


Influenza

Influenza is a viral infection of the lungs and upper respiratory tract.

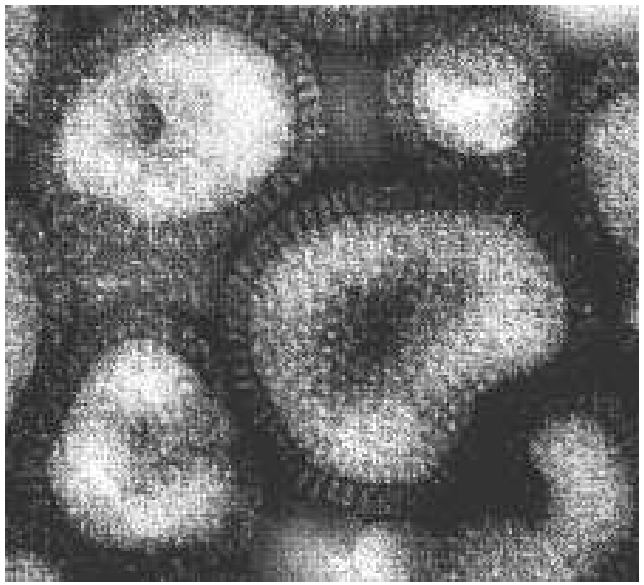
Three basic types: Influenza A, B, and C.

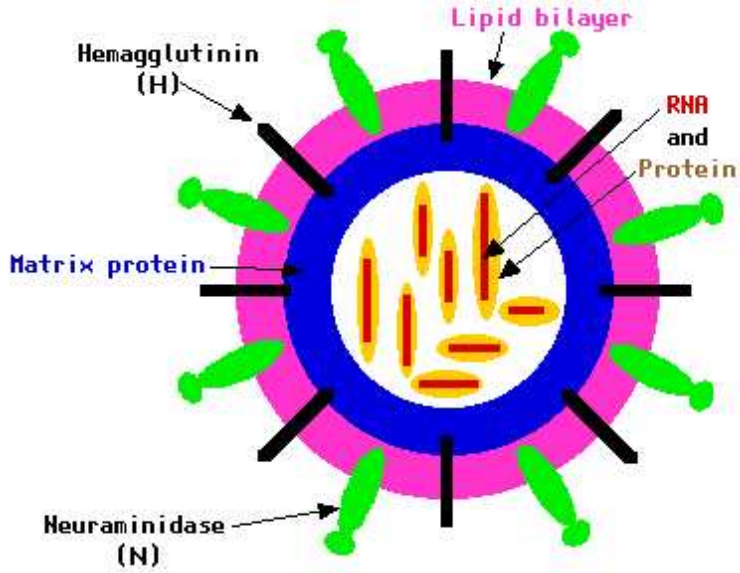
Influenza A is the type responsible for regular outbreaks and global pandemics.



Influenza A Virus Replication

Influenza A



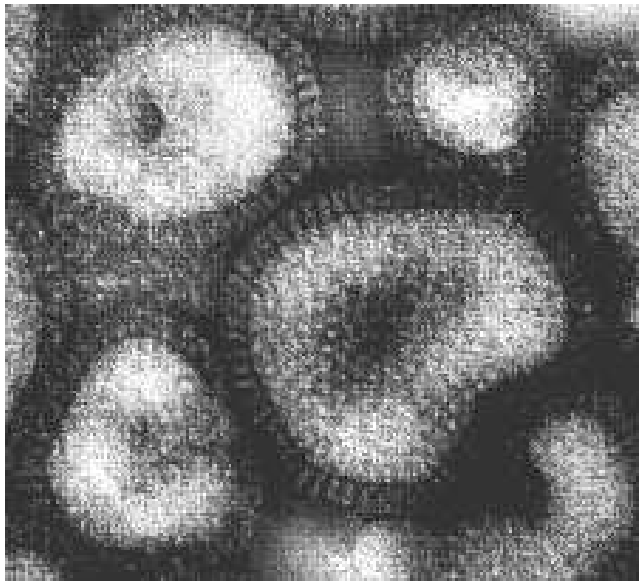


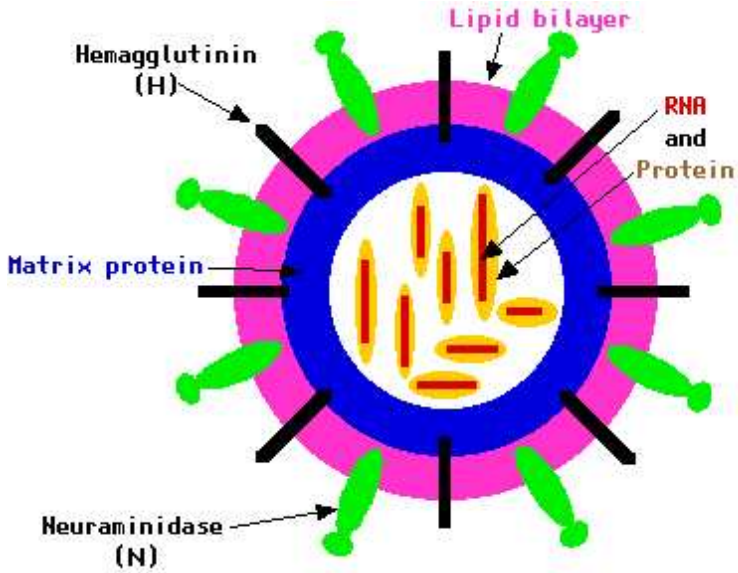
- Virus's genetic material contains eight genes encoding for eleven proteins.
- The hemagluttinin and neuraminidase act as antigens that are recognized by the immune system.

  There are 16 known forms of hemaggluttinin and 9 known forms of neuraminidase.

  The form of each is used to classify the viruses. For example, an H5N1 virus has the fifth known form of hemaggluttinin and the first form of neuraminidase.

Influenza A (cont'd)





- Until recently, H1, H2, H3 and N1 and N2 are the only forms found in human infections.

  Recent avian flu viruses are H5N1 and have infected humans, though for the moment they do not transmit rapidly.

- This virus is capable of rapid genetic change in mammals. It's genetic complex lacks "proofreading" capability, so almost 20% of virions produced will contain a mutation.

Mutations

Mutations are *random* changes in an organism's genetic material (DNA or RNA).

Many mutations are caused by copying errors during replication and some by induced chemical change (e.g., by radiation).

In a broader sense, mutations can arise through a variety of effects and interactions, including:

- Translocation and recombination
- Lateral transfer mutations (plasmids)
- Viruses
- Selfish genetic elements (e.g., transposons)

Built-in mechanisms to repair DNA are fast and very accurate, but as with anything, they are not perfect.

Types of Mutations

Substitution: ACGTG
AGGTG

Insertion: TAAGCG
TACCGAGCG

Deletion: TAAGCG
TACG

Frame shift: ACG TTA AAG GCG C...
CGT TAA AGG CGC

Duplication: AATCCG
AAAAATCCG

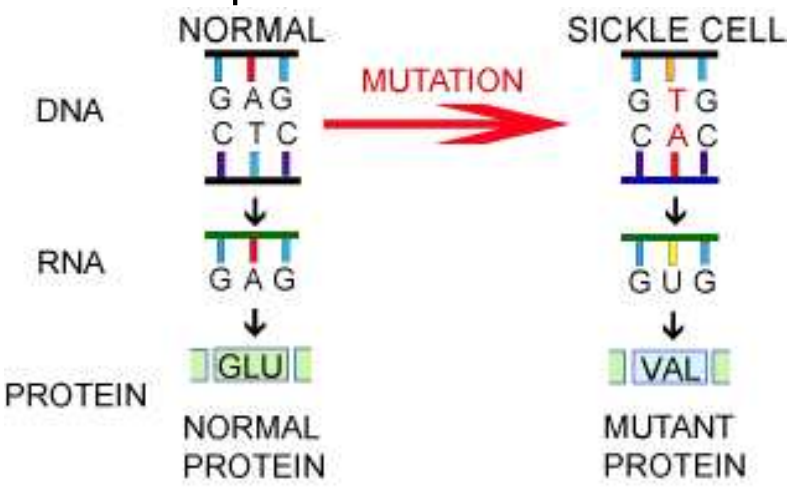
...

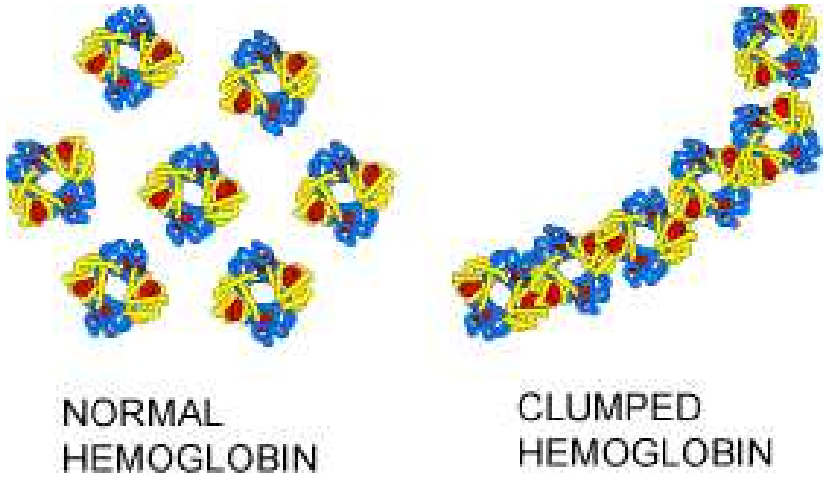
More Detail: Substitution

Substitution can lead to a variety of changes.

- Mis-sense mutation: the changed nucleotide alters the codon to produce a different amino acid in a viable protein.

  Example 1: Sickle-cell Anemia





  Example 2: Cystic Fibrosis

  Defects in a protein (cystic fibrosis transmembrane conductance regulator) cause symptoms of the disease, but no one type of mutation is responsible.

  At least one form of mutation that causes the disease is a mis-sense substitution.

More Detail: Substitution

- Non-sense mutation: the changed nucleotide creates a nonviable gene product. For instance, if a codon is changed to a STOP codon, the truncated polypeptide will likely not function properly.

  Example: Cystic Fibrosis, a different mutation

- Silent mutation: the changed nucleotide does not change the amino acid produced, and there is no phenotypic effect.
- Splice-site mutation: change to the nucleotides that are cut out during gene editing (introns). Can affect the final RNA molecule that is produced.

Key points about Mutations

- Mutations can be harmful, neutral, or beneficial.
- Mutations can have large, small, or no effect on the phenotype

  Examples: homeotic mutations, DDT-resistance, lethals, ...
- Mutations are rare.
- Only germline mutations (as opposed to somatic mutations) are heritable and thus of evolutionary import.
- Male's tend to have more mutations than females, but not as many more as might be predicted.

Mutation Rates

The human genome has 3 billion base pairs. Even considering only substitutions, each cell has 6 billion chances for a mutation.

Still, the copy/repair process is amazingly accurate. One estimate states that there is on average about 1 uncorrected error for every 50 million nucleotides copied.

But that's still about 100 mutations per cell. That seems high?

Ameliorating factors:

- Much of the genome is non-coding DNA or is spliced out of the gene before expression.
- Redundancy in the codon set
- Dominance can suppress recessive mutations. (Compensation.)

Estimating Mutation Rates

What makes this difficult?

Why might it be useful?

Is the frequency of a mutation a good indicator of the rate?

Recent work with *C. elegans* produce sharp estimates of mutation rates by carefully managing the environment over many (several hundred) generations.

Are mutations directional?

Leading Example: bacterial resistance spreads through the population in response to antibiotics (esp. their incomplete use).

How does this happen?

Are mutations directional?

Leading Example: bacterial resistance spreads through the population in response to antibiotics (esp. their incomplete use).

Two hypotheses

1. The genes conferring resistance existed in the population but became favored by selection when the antibiotic is used.
2. The presence of the antibiotic led to the evolution of the resistance genes.

Can we distinguish these empirically?

Are mutations directional? (cont'd)

Famous experiment (Lederberg and Lederberg, 1952):

- Generated colonies of bacteria from a single clone.
- Cultured them for some time on plates.
- Transfered part of colonies (with felt) to an identical plate, producing two plates with identical groups of colonies.

- Exposed one plate to an antibiotic and noted which colonies survived.
- Examined corresponding colonies on other plate.
- All of the latter bacteria were resistant even though they had never been exposed to it.

Back to Influenza A

- Pandemics and epidemics

Date	Subtype	Notes
1918	H1N1	pandemic ("Spanish" flu), devastating
1957	H2N2	pandemic ("Asian" flu)
1962	H2N2	epidemic
1964	H2N2	epidemic
1968	H3N2	pandemic ("Hong Kong" flu),
1976	H1N1	swine flu in recruits, no epidemic
200?	H5N1	avian flu in humans?? pandemic???

  A pandemic strain must be "new" to its host and at the same time well adapted to it.

- Mutations and reassortment of genetic strands
- Antigenic shift and Antigenic drift
- Phylogenetics