A quasispecies

is a population of RNA or DNA molecules:

ATCAGGACTCA
ATCGGGACTCA
ATCAGGAATCA
....

Manfred Eigen & Peter Schuster 1977

The Quasispecies equation

Frequency of type j

Mutation from j to i

$$\dot{x}_i = \sum_{j=1}^{n} x_j f_j Q_{ji} - f(\vec{x}) x_i$$

Fitness of type j

Average fitness

$$f(\vec{x}) = \sum_i f_i \ x_i$$

$$\sum_i x_i = 1$$

The Quasispecies equation

$$\dot{x}_i = \sum_{j=1}^{n} x_j f_j Q_{ji}$$

$$\dot{x}_i = \sum_{j=1}^{n} x_j Q_{ji} + f_i x_i$$

Mutation matrix

$$Q_{ij} = p^{H_{ij}}(1-p)^{L-H_{ij}}$$

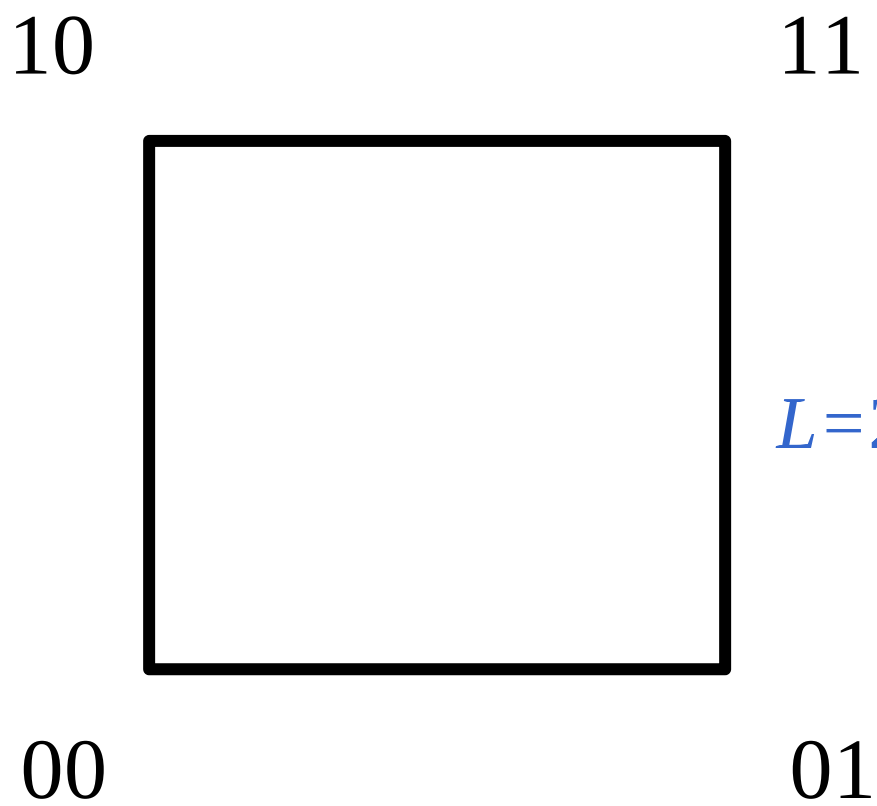
p...mutation rate per bit

L...genome length

H_{ij}...Hamming distance

(= number of point mutations
between i and j)

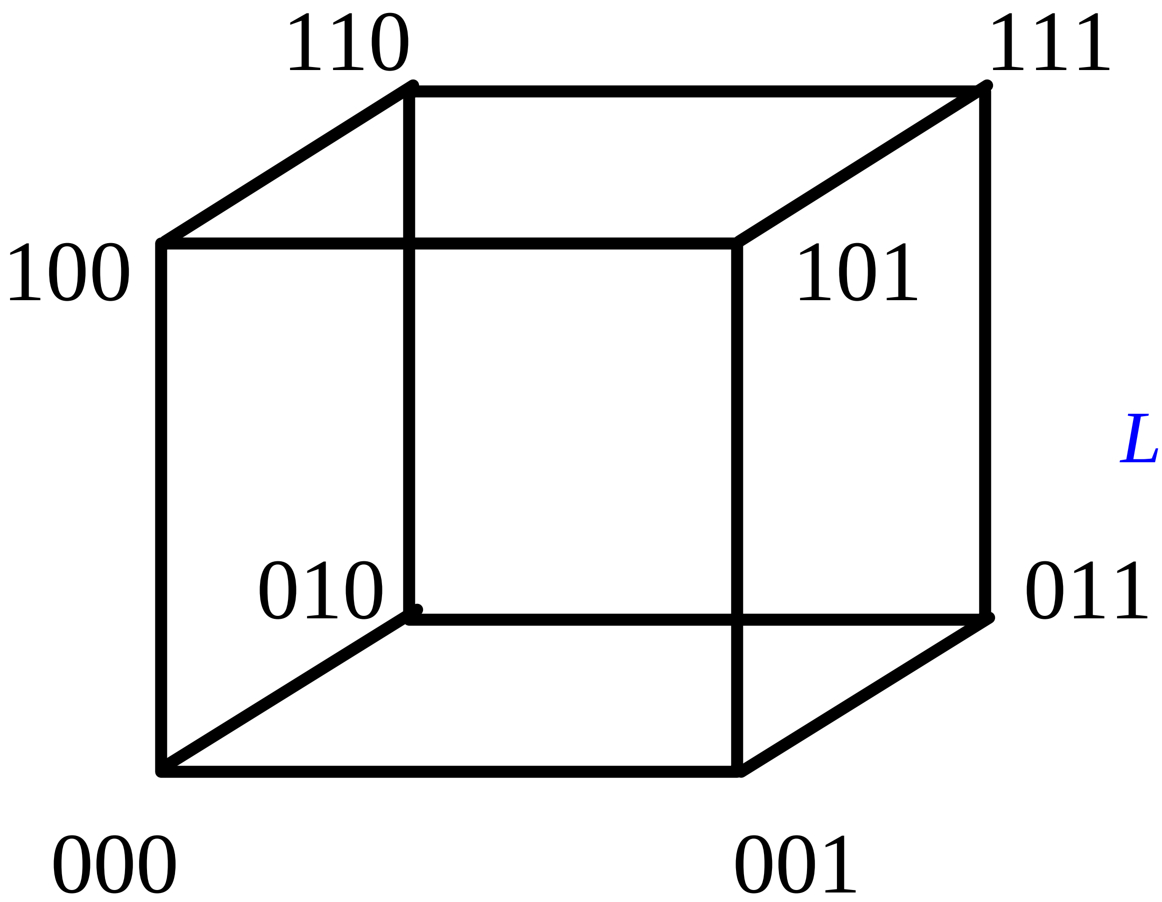
Sequence space



L=2 dimensions

Hamming (or Manhattan) metric

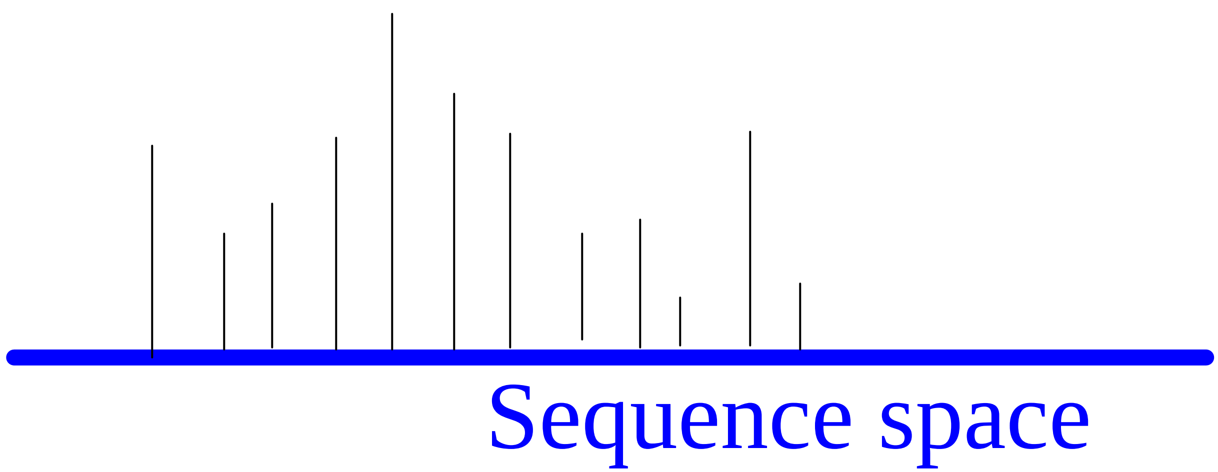
Sequence space



L=3 dimensions

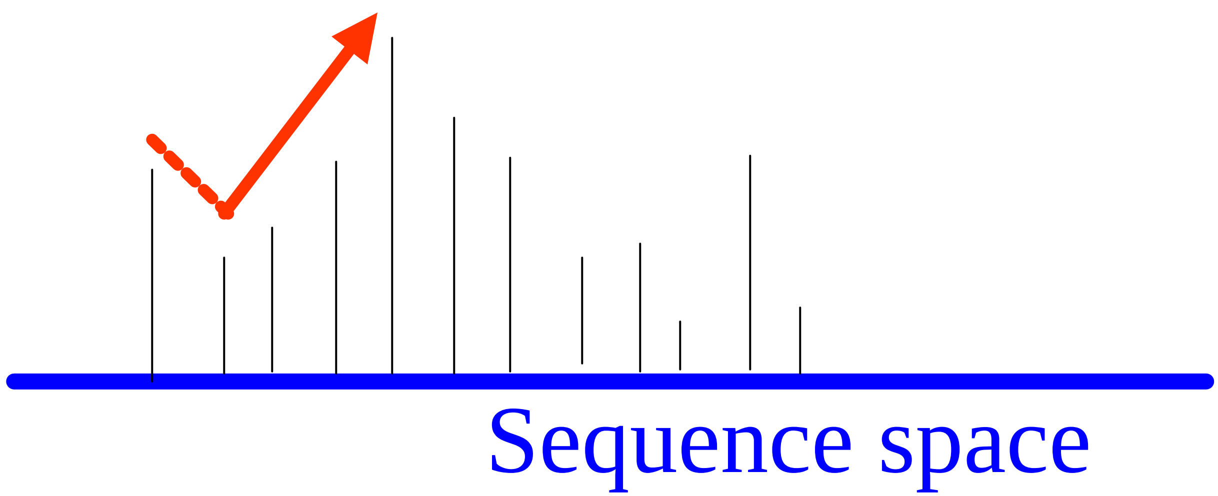
Fitness landscape

Each point in sequence space is assigned a number.



Evolution

… is adaptation of the quasispecies on the fitness landscape.

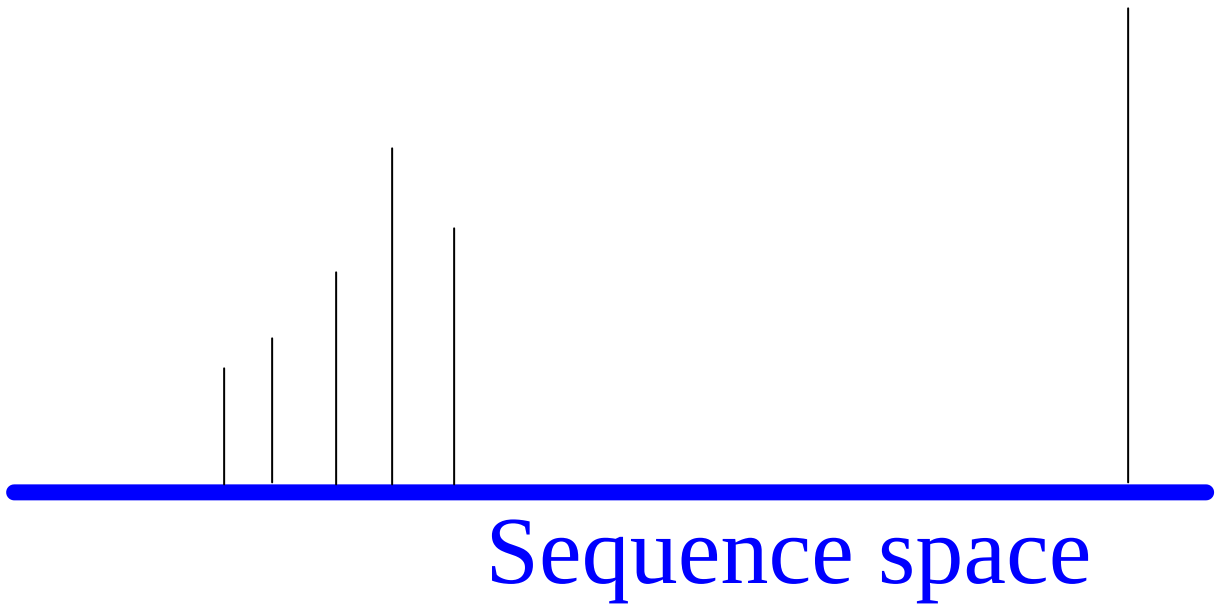


Search process

Not a single sequence, but the a whole ensemble of sequences explores the fitness landscape.

Survival of the quasispecies

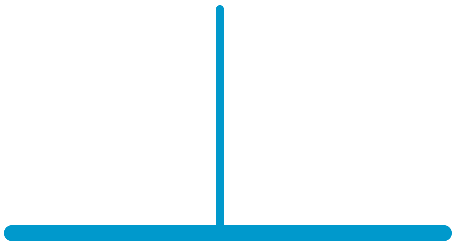
Not the fittest sequence (maximum *a*), but the fittest quasispecies is selected (maximum *Phi*).



Error threshold

If the mutation rate is to high,
the quasispecies cannot maintain
genetic information.
Advantageous mutants disappear.

Error threshold



Wildtype $\dot{x}_1 = f_1 x_1 Q - x_1 f$

Mutants $\dot{x}_2 = f_1 x_1 (1-Q) + f_2 x_2 - x_2 f$

Wildtype survives if $f_1 Q > f_2$

Error threshold implies a maximum genome length

$$L < 1/p$$

genome length

mutation rate per base

Error threshold

Adaptation is only possible if the mutation rate is below a certain threshold.

$$p < 1/L$$

Localization

Applications of quasispecies theory

- origin of life
- in vitro evolution
- viruses
- bacteria
- cancer

Extensions

- More complicated mutation operator (insertion, deletion, recombination, etc)
- Stochastic formulation (finite population)