Predictive Modeling of Evolution
Influenza Virus
The Influenza Genome

- HA hemagglutinin
- NA neuraminidase
- M2 proton channel
- M1 matrix protein
- Lipid bilayer
- PA RNA polymerase complex
- PB1 RNA polymerase complex
- PB2 RNA polymerase complex
- RNA + NP Nucleoprotein
- PB2 1, 2, 3
- PB1 4, 5
- PA 6, 7
- HA 8
- NP 9
The “Red Queen Race” of Influenza and the Human Immune System

- Co-evolutionary process of binding (AA-AA interactions)
- Strong non-equilibrium dynamics, with highly time dependent fitness

Static Fitness Landscape (Equilibrium)

Sea-scape (Non-Equilibrium)
The “Red Queen Race” of Influenza and the Human Immune System

"The Red Queen has to run faster and faster in order to keep still where she is. That is exactly what you all are doing!"
The need to predict

- Vaccine production and development requires 6-8 months
- The choice of the strain to be used in the vaccine has to be made almost 1 year in advance!
Genetic signature of the flu evolution: Phylodynamics

- Hemagglutinin Gene of the H3N2 Flu virus
- Strain isolated in 1 year = 1 coalescent tree (one cluster in the full tree)
- Signature of rapid evolution (Trunk of the tree)
Genetic signature of the flu evolution: Phylodynamics

- Neutral Mutations (Synonymous)
- Non-Epitope Mutations
- Epitope Mutations

Antibody
Genetic signature of the flu evolution: Phylodynamics

(a) Neutral Mutations (Synonymous)
(b) Non-Epitope Mutations
(c) Epitope Mutations
Fitness Model
Fitness model

- Evolutionary Units: phylogenetic clades
- Clade evolutionary dynamics
- Fitness:

\[
\hat{X}_v(t+1) = \sum_{i:v,t} x_i \exp(f_i)
\]

\[
f_i = f_0 - \mathcal{L}(a_i) - \sum_{j: t_j < t_i} x_j C(a_i,a_j)
\]

Non epitope Mutations

Epitope Mutations
Fitness model

- Fitness:
  \[ f_i = f_0 - \mathcal{L}(a_i) - \sum_{j: t_j < t_i} x_j C(a_i, a_j) \]

- Stability
- Antigenicity
Fitness model

- Model parameters inferred with a Maximum Likelihood inference, from the phylogenetic tree.

- Model prediction tested with a retrospective analysis:

  - Predicted Ratio
    \[ \hat{W}_v = \hat{X}_v(t+1)/X_v(t) \]

  - Posterior Ratio
    \[ W_v = X_v(t+1)/X_v(t) \]
Fitness model

(a) Predicted ratio, $\hat{W}$ vs. Posterior ratio, $W$

(b) Predicted substitutions

(c) Posterior substitutions

Legend:
- Nonsynonymous, in epitopes
- Nonsynonymous, outside epitopes
- Synonymous

$\hat{X}(t+1)$

Predicted frequency change

$< -0.1$ to $> 0.2$
Fitness model

- **Vaccine Choice:**
  Maximize cross-immunity amplitude

\[ C_v(t+1) = \sum_{i:t} x_i \exp(f_i)C(a_i,a_v) \]

- Comparison with the actual cross-immunity amplitude of actual vaccine
Branching Index
Travelling fitness wave

- Traveling fitness model
  - High mutation rate of beneficial mutations
  - High selective pressure
  - Non-equilibrium model (can be extended to equilibrium)
Traveling fitness model

- High mutation rate of beneficial mutations
- High selective pressure
- Non-equilibrium model (can be extended to equilibrium)
Travelling fitness wave: a new universality class

HA tree (1 year)  Traveling wave tree  Kingman tree

Frequency spectrum
Travelling fitness wave: a new universality class

HA tree (1 year)  Traveling wave tree (Bolthaus Snitman)  Kingman tree

Frequency spectrum

Realistic (?)
Branching Index

Local density of the branches around a node → Fitness of a node

Branching Index
(integrated exponentially discounted tree length surrounding a node)

\[
\lambda_i(\tau) = m_{ij} + \sum_j m_{ij}\
\]

\[
m_{ij} = \tau \left(1 - e^{-b_i/\tau}\right) + e^{-b_i/\tau} \sum_j m_{ij}\
\]

\[
m_{ij} = \tau \left(1 - e^{-b_{ij}/\tau}\right) + e^{-b_{ij}/\tau} \left[m_{ij} + \sum_{k \neq j} m_{ik}\right]\
\]
Branching Index

Local density of the branches around a node

Fitness of a node

Tested with simulated data
Branching Index

- Model prediction tested with a retrospective analysis:

\[ d = \frac{\Delta(\text{prediction}) - \Delta(\text{minimal})}{1 - \Delta(\text{minimal})} \]

- \( d = 0 \): Optimal Choice
- \( d = 1 \): Random Choice
Predicting Evolution?
Predicting Evolution?

![Graphs showing evolutionary trends](image)

**Part a:** Yeast *S. cerevisiae*

**Part b:** Human influenza

**Part c:** Blood cancer

**Part d:** Immune repertoire

**Part e:** Schematic representation of evolutionary changes