Origin of directional epistasis in RNA folding

Claus O. Wilke¹, Richard E. Lenski², and Christoph Adami¹

¹Digital Life Laboratory 136-93
California Institute of Technology
Pasadena, CA 91125

²Center for Biological Modeling
Michigan State University
East Lansing, MI 48824

Decay Functions

- Investigate fitness decay functions \( w(d) \).
- Decay functions give change in average fitness with genetic distance \( d \) from some reference sequence.
- In RNA case: \( w(d) \) is the number of neutral folds divided by the total number of sequences at distance \( d \).

Fitness Landscape

Null hypothesis (no epistasis):
\[ w(d) \approx \exp(-\alpha d) \]

With epistasis:
\[ w(d) \approx \exp(-\alpha d^\beta) \]
\( \beta > 1 \): synergism
\( \beta < 1 \): antagonism

Exponential decay is caused by the tre-like structure of the genotype space (circles: viable sequences, crosses: inviable sequences).

Correlation between \( \alpha \) and \( \beta \)

Measured \( \alpha \) and \( \beta \) for 100 random RNA sequences of length 76 (RNA folding done with Vienna package).

Result:
\( \alpha \) and \( \beta \) are strongly correlated.

\( \rightarrow \) Sequences for which single mutations have a large effect are strongly antagonistic, sequences for which single mutations have a small effect are only slightly antagonistic or even synergistic.
Origin of correlation

high $\alpha$ $\rightarrow$
high density of neutral sequences $\rightarrow$
density decreases as $d$ increases $\rightarrow$ high $\beta$

likewise:
low $\alpha$ $\rightarrow$ low $\beta$.

Origin of shift in $\beta$

Hypothesis: background of compensatory mutations shifts distribution of $\beta$.

Here, compensatory mutations are defined as those that lead from an inviable sequence to a viable one that is not part of the original neutral network.

Knowledge of neutrality at distance $d$ allows us to disentangle compensatory mutations from other mutations.

Distribution of Density Changes

The parameter $m$ measures the change in neutrality with distance $d$ from the reference sequence.

Why is $\beta$ always $\leq 1$, when neutrality changes ($m$) go both ways?
**Adjusted $\alpha$ and $\beta$**

The adjustment results in an upwards shift of $\beta$.

The epistasis parameter $\alpha$ is almost unaltered after the adjustment.

**Distribution of $\beta$**

We use two different methods of adjustment.

In both cases, a significant percentage of cases has $\beta > 1$.

(Measured: $\langle \beta \rangle = 0.89 \pm 0.01$, Adjusted M1: $\langle \beta \rangle = 1.05 \pm 0.01$, Adjusted M2: $\langle \beta \rangle = 0.98 \pm 0.01$)

**Conclusions**

- Strength and direction of epistasis depend on position of chosen reference sequence in genotype space.
- Synergistic epistasis: reference sequence is in center of a dense cluster of viable sequences.
- Antagonistic epistasis: reference sequence is outside of a dense cluster of viable sequences.
- Background of compensatory mutations leads to excess of antagonistic epistasis.

**References**


Origin of directional epistasis 
in RNA folding

Claus O. Wilke\textsuperscript{1}, Richard E. Lenski\textsuperscript{2}, and Christoph Adami\textsuperscript{1}

\textsuperscript{1}Digital Life Laboratory 136-93
California Institute of Technology
Pasadena, CA 91125

\textsuperscript{2}Center for Biological Modeling
Michigan State University
East Lansing, MI 48824

Decay Functions

- Investigate fitness decay functions \( w(d) \).

- Decay functions give change in average fitness with genetic distance \( d \) from some reference sequence.

- In RNA case: \( w(d) \) is the number of neutral folds divided by the total number of sequences at distance \( d \).

Fitness Landscape

Null hypothesis (no epistasis):
\[ w(d) \approx \exp(-\alpha d) \]

With epistasis:
\[ w(d) \approx \exp(-\alpha d^\beta) \]
\( \beta > 1 \): synergism
\( \beta < 1 \): antagonism

Exponential decay is caused by the tre-like structure of the genotype space (circles: viable sequences, crosses: inviable sequences).

Correlation between \( \alpha \) and \( \beta \)

Measured \( \alpha \) and \( \beta \) for 100 random RNA sequences of length 76 (RNA folding done with Vienna package).

Result:
\( \alpha \) and \( \beta \) are strongly correlated.

\[ \rightarrow \text{Sequences for which single mutations have a large effect are strongly antagonistic, sequences for which single mutations have a small effect are only slightly antagonistic or even synergistic.} \]
Origin of correlation

high $\alpha \rightarrow$
high density of neutral sequences $\rightarrow$
density decreases as $d$ increases $\rightarrow$ high $\beta$
likewise:
low $\alpha \rightarrow$ low $\beta$.

Origin of shift in $\beta$

Hypothesis: background of compensatory mutations shifts distribution of $\beta$.

Here, compensatory mutations are defined as those that lead from an inviable sequence to a viable one that is not part of the original neutral network.

Knowledge of neutrality at distance $d$ allows us to disentangle compensatory mutations from other mutations.

Distribution of Density Changes

The parameter $m$ measures the change in neutrality with distance $d$ from the reference sequence.

Adjusting for Compensatory Mutations

We can separate $w(d)$ (solid line) into contributions from original neutral network (dashed line) and contribution from compensatory mutations (dot-dashed line).

Why is $\beta$ always $\leq 1$, when neutrality changes ($m$) go both ways?
**Adjusted $\alpha$ and $\beta$**

The adjustment results in an upwards shift of $\beta$.

The epistasis parameter $\alpha$ is almost unaltered after the adjustment.

**Conclusions**

- Strength and direction of epistasis depend on position of chosen reference sequence in genotype space.
- Synergistic epistasis: reference sequence is in center of a dense cluster of viable sequences.
- Antagonistic epistasis: reference sequence is outside of a dense cluster of viable sequences.
- Background of compensatory mutations leads to excess of antagonistic epistasis.

**Distribution of $\beta$**

We use two different methods of adjustment.

In both cases, a significant percentage of cases has $\beta > 1$.

(Measured: $\langle \beta \rangle = 0.89 \pm 0.01$, Adjusted M1: $\langle \beta \rangle = 1.05 \pm 0.01$, Adjusted M2: $\langle \beta \rangle = 0.98 \pm 0.01$)

**References**
