Evolution of coding structure
RNA world: sequence - structure - interactions
2 images of RNA world

individual complexity

ecosystem complexity

sequence to structure
replicator to wave/vesicle

"RNA ideal evolvable molecule"
the RNA world
Today: RNA in space

Themes

Structured based modeling

Individual and/or ecosystem based complexity
ecosystem diversification and mutation rate

Evolution of coding structures (cont)
multiple coding
mutational neighborhood

RNA even more evolvable than seen so far
RNA world: Preconceived networks vs evolving individuals, emerging species, emerging interactions

- **Structured individuals**
  here RNA sequences (+ and - strands)
  if folding in predefined structure: replicase

- **No predefined target or fitness**

- **No predefined interactions**
  but predefined reactions

**DO SPECIES/ INTERACTION NETWORKS EVOLVE?**
**DOES EVOSYSTEM COMPLEXITY EVOLVE?**
genotype - phenotype - ecosystem mapping

Genotype → Phenotype → Interaction

sequence

structure

ecosystem

cat-C
par-G

cat-A
par-U

feedback from higher levels to lower levels in evolving system

interacting RNA’s

Complex formation happens 5’-end $\rightarrow$ 3’-end

(“strong” altruism)

1. $X + Y \xrightleftharpoons[k_1]{k_{-1}} C_{X\sim Y}$ or $\xrightleftharpoons[k_2]{k_{-2}} C_{Y\sim X}$
2. $C_{X\sim Y} + \Phi \xrightarrow[\kappa]{\kappa} X + Y + Y^{-1}$
3. $X \xrightarrow[d]{d} \Phi$

only structure + reaction
no fitness function and no interaction predefined
Maximum mutation rates ($\mu = .015$): is only below information threshold for evolved coding structure ONE quasispecies

initial population $\quad$ dynamics with mutation

$= \text{after stopping mutation}$
High mutation rate ($\mu = 0.015$)

Population structure of $+$ strands

- Phylogeny reveals patterns in population of genotypes

- No clade patterns

- Population is supported by various genotypes

- One quasi-species

<table>
<thead>
<tr>
<th>Color</th>
<th>Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyan</td>
<td>Catalyst</td>
</tr>
<tr>
<td>Red</td>
<td>Non-catalyst</td>
</tr>
</tbody>
</table>
High mutation rates ($\mu = .015$) sequence structure: symmetry breaking: only $+$ strands catalytic

- Very high C frequency in 5’-end
- High G frequency in 3’-end
  - $\rightarrow$ many GC pairs
- Many interspersing U in 3’-end
  - $\rightarrow$ prevents base-pair formation *in homo*
- No 5’-end in template strand
  - $\rightarrow$ prevents non-functional complex formation
Sequence is delicately tuned up
- Almost all base-pairs are GC
- Many other G and C that should not pair
  - Difficult to form correct base-pairs
- High sequence conservation in all positions
  - Loop region must be tuned too
lowering mutation rates \((\mu - .13)\) : SPECIATION

\[ m = .014 \quad \rightarrow \quad m = .013 \]
lowering mutation rates: \((\mu = 0.13)\) population structure

- Two quasi-species
  - distinct sequence classes
  - Catalyst & Non-catalyst

<table>
<thead>
<tr>
<th>Color</th>
<th>Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyan</td>
<td>Catalyst</td>
</tr>
<tr>
<td>Red</td>
<td>Non-catalyst</td>
</tr>
</tbody>
</table>
lowering mutation rates: \( (\mu = 0.13) \) sequence structure

- **Catalyst**
- **Non-catalyst → Parasite**

- No catalytic structure in both strands
- Long 3’-end with many G
- No 5’-end in both strands
- Sequence conservation patterns & secondary structure
Parasite invades in periphery of QS

- Population of Sequences
- Genotype & Phenotype
- Space & Time

Hamm. dist. from master sequence
Lower mutation rate $\mu = .008$: 3 quasispecies

Population of Sequences

Genotype & Phenotype

Space & Time

A catalyst: HIGH neutrality (ca 50%)
Lower mutation rate $\mu = 0.004$: 4 quasispecies
evolved 4 species system; evolved interaction topology

ECOsystem ($\mu = 0$)

EVOL. system ($\mu = .004$)

Direct Interaction structure

<table>
<thead>
<tr>
<th></th>
<th>C-catalyst</th>
<th>A-catalyst</th>
<th>G-parasite</th>
<th>U-parasite</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CYAN</td>
<td>MAGENTA</td>
<td>RED</td>
<td>GREEN</td>
</tr>
<tr>
<td></td>
<td>cat. str.</td>
<td>comp.</td>
<td>cat. str.</td>
<td>comp.</td>
</tr>
<tr>
<td>C-cat</td>
<td>0.52</td>
<td>0.87</td>
<td>0.36</td>
<td>0.45</td>
</tr>
<tr>
<td>A-cat</td>
<td>0.39</td>
<td>0.05</td>
<td>0.50</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.81</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.14</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.55</td>
</tr>
</tbody>
</table>
From Coding structure to ecosystem based information accumulation
Conclusion

Very stable multi-(quasi)species systems evolves

Interaction topology different from anything studied before.

Variability increases with decreasing mutation rate speciation

Ecosystem based "solution" only at lower mutation rates

EVOLVED genotype-phenotype-interaction-spatial structure mutual dependent (and "make sense" in relation to each other)

Evolved, niche dependent mutational landscape
Evolution of coding structure at high mutation rates
Mutational neighborhood
High mutation rate: 1 quasispecies LOW variability

**mutational neighborhood of master seq.: STEEP**

EVL (evolved) | optimal repl | av. random
---|---|---
black replicator; blue rest

Colizzi & H. 2014
High mut. rate: 1 quasispecies LOW variability

mutational NB: STEEP and “special”

EVALVED  optimal repl  av. random

black replicator; blue rest; yellow parasites
1 quasispecies: codes for multiple functions

Mutational NB: STEEP and “special”

Evolved optimal repl av. random

Black repl.; blue rest; yellow parasites; green helpers; red stallers; gray junk
mutational neighborhood at larger Hamming distances

Top follow replicases with \( \geq \) replic rates masterseq. bottom follow replicases with \( < \) replic rates masterseq.
quasispecies composition in field

weakly reflects mutational neighbourhood
more replicators (because of replication,
less helpers, more stallers (like neighborhood of other replicases)

Replicases with 'good' mutational neighborhood overrepresented.
Helpers “help”
change in junk $\rightarrow$ extinction
change in empty $\rightarrow$ extinction

in simplified ODE model:
increases max $\mu$ without parasites
decreases max $\mu$ with parasites
Stallers “stall”
change in junk -> increases density
BUT master seq. replaced
’pseudo stallers’ evolve
to change into empty space
parasite lineage evolves!

in simplified ODE model:
protects against parasites

\[
\begin{align*}
2X & \xrightarrow{a_{xx}} C_{xx} \\
\xrightarrow{b_{xx}} & 3X \\
\kappa,\theta, (1-\mu) & \rightarrow 2X + S \\
\kappa,\theta, \mu \cdot \lambda_S & \rightarrow 2X + P \\
\kappa,\theta, \mu \cdot \lambda_P & \rightarrow 2X + J \\
\kappa,\theta, \mu \cdot \lambda_J & \rightarrow 2P + X \\
2P + X & \xrightarrow{a_{px}} C_{px} \\
\xrightarrow{b_{px}} & \rightarrow P + X + S \\
\kappa,\theta, (1-\mu) & \rightarrow P + X + J \\
X + S & \xrightarrow{a_{xs}} C_{xs} \\
\xrightarrow{b_{xs}} & \rightarrow P + S + C_{ps} \\
\kappa,\theta, \mu \cdot \lambda_S & \rightarrow X, S, P, J, C_{xx}, C_{px}, C_{xs}, C_{ps} d \rightarrow \theta,
\end{align*}
\]

with parasites: x-axis: fraction staller-mutants
## Variability of evolved quasispecies

<table>
<thead>
<tr>
<th>mut. nei.</th>
<th>$\mu_{max}$</th>
<th>replic. rates</th>
<th>competition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+/+</td>
<td>+/-</td>
<td>7</td>
</tr>
</tbody>
</table>

### Steep quasispecies

1. $0.0164$  
   - 0.902 0.858  
   - 0.914 0.831  
   - ✔ ✔

2. $0.0154$  
   - 1.000 0.878  
   - 0.932 0.854  
   - ✔ ✔

3. $0.0151$  
   - 1.000 0.777  
   - 0.870 0.744  
   - ✔ ✔

4. $0.0145$  
   - 1.000 0.817  
   - 0.866 0.777  
   - ✔ ✔

5. $0.0151$  
   - 1.000 0.777  
   - 0.818 0.731  
   - ✔ ✔

6. $0.0143$  
   - 1.000 0.777  
   - 0.858 0.729  
   - × ×

### Flat quasispecies

7. $0.0154$  
   - 0.725 0.817  
   - 0.892 0.798  

8. $0.0149$  
   - 0.902 0.817  
   - 0.872 0.792  

### Colors
- Random colors
- Majority function
- Distance from masterseq
Mutational neighborhood of 2 functionally equivalent RNA’s

(c) $\mu = 0.008$

- S,

black replicator; yellow parasite; green helper; red staller

Conclusions RNA world at high mutation rates

Evolution of very specific coding structure.

One mastersequence codes for functional diverse ecosystem

Decoded by mutations (hence clearest at high mutation rates)

In steep quasispecies most pronounced (best 'control')

Individually coded but ecosystem based diversity evolves and persists close to the Information Threshold