High rates of duplications and deletions prevent evolutionary deterioration

The role of transcriptional load-associated mutations for genome stability and fitness

Enrico Sandro Colizzi and Paulien Hogeweg

Utrecht University

January 16, 2017
rDNA copy number is controlled in all living beings:

# copies in bacteria: 1 → 16
# copies in eukaryotes: 150 → $10^4$
Bacteria:
rDNA copy number reflects ecological strategies

growth vs. efficiency trade-off

Roller et al. 2016
However:

- if you believe the regression
- many bacteria can switch fast ↔ efficient growth
- rRna at origin of replication can effectively increase copy number during fast growth
Eukaryotes:

<table>
<thead>
<tr>
<th>Species</th>
<th>Copy no.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saccharomyces cerevisiae</td>
<td>150</td>
</tr>
<tr>
<td>Drosophila melanogaster</td>
<td>240</td>
</tr>
<tr>
<td>Xenopus laevis</td>
<td>600</td>
</tr>
<tr>
<td>Homo sapiens</td>
<td>350</td>
</tr>
<tr>
<td>Arabidopsis thaliana</td>
<td>570</td>
</tr>
<tr>
<td>Pisum sativum (pea)</td>
<td>3,900</td>
</tr>
<tr>
<td>Zea mays (maize)</td>
<td>12,000</td>
</tr>
</tbody>
</table>

Long and Dawid 1980, Kobaiaishi 2011
Yeast rRNA gene cluster: ~ 150 copies

No difference in growth rate with different copy numbers, as long as > 30

similar number to bacteria
Variation in copy number occurs often, also due to collisions between transcription and replication complexes. Conflicts between transcription and replication can lead to genome and epigenome instability.

French et al. 2003, Takeuchi et al. 2003
Few rRNA gene copies ($\sim 30$)+ caloric excess (via TOR pathway)

$\downarrow$

copy number increases until normal again

Jack et al. 2015
WHY?

No growth difference $\Rightarrow$ No direct selection

Second order effect?
Long term genome stability

We make a model to understand the effect of mutations observed in yeast
Cells grow and divide:

Growth: minimal metabolism and regulation to make macromolecules

Division: target volume $\propto (\text{genome size})^{0.9}$
Volume = sum of macromolecules

Macromolecules and gene types:
- T: Enzymes
- Q: Housekeeping proteins
- Rp: Ribosomal protein
- Rr: Ribosomal RNA
- Background mutations: faulty replication
- Transcription-induced mutations: Conflicts transcription-replication mutations $\propto$ transcription load
- Phenotypic mutations in regulation parameters
- Inactive genes retain promoters (still transcribed)

Selection for faster growth provided homoeostasis
Equal proportion of mutations

<table>
<thead>
<tr>
<th>Genome size</th>
<th>Active genes</th>
<th>Inactive genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{BG}$</td>
<td>$\mu_{BG+TR}$</td>
<td>$\mu_{BG}$</td>
</tr>
</tbody>
</table>

Interdivision time

$\mu_{BG} = 0.084$

$\mu_{BG+TR}$

$max(\mu_{TR}) = 0.074$

1/3 duplications,
1/3 deletions,
1/3 inactivations

Large mutational load: WHY?
Ancestors:
more active genes than rest of the population

Short term fitness = more active genes
Evolution with constant regulation: early vs late parameters

- Short-term beneficial mutations
  \[ \downarrow \]
- Long-term accumulation of deleterious effects
Increase frequency of dupl and del $\rightarrow$ higher fitness and genome (and higher tr. load)

**Genome size**

**Active genes**

**Inactive genes**

Inter-division time

transcr mut.

$\mu_{BG}$: background mutations only

$\mu_{TR (1,2,3)}$: background mutations + transcriptional mut.

dupl., del., inact.

$\mu_{TR 1}$: 1/3, 1/3, 1/3,

$\mu_{TR 1}$: 5/12, 5/12, 2/12,

$\mu_{TR 3}$: 1/2, 1/2, 0
Evolution avoids mutational runaway process
Slightly deleterious mutations do not accumulate

Short-term “larger” effect of mutations
- Duplications of active genes + Deletion of inactive genes $\rightarrow$ positively selected
- Smaller relative effect of inactivation despite being the same in absolute terms
A simpler model:

One gene type: active (a) or inactive (i)

Genome size (G) = active + inactive genes

Target gene copy number $\alpha$, e.g. 100 copies

fitness = $(a/G) \times (1 - |1 - a/\alpha|)$
Lower mutation rates ($\mu = 0.003$) equal proportion dupl/del/inact.
Higher mutation rates ($\mu = 0.011$)
Biased dupl/del/inact. (5/11,5/11,1/11)
Recombination does not solve *this* problem of accumulation of slightly deleterious mutations.
Simpler model:

- NOT a good “first model”: too artificial
- good “model of a model”: reproduces results and easy to test hypothesis
- (fast)
When proportion of transcription-induced mutations can mutate $\rightarrow$ bias towards transcription-induced duplications and deletions is selected.

Robust results for range of values of background inactivations.
Yeast-like conditions: higher background rate of inactivations and deletions. Proportions of transcription-induced mutations evolve.

\[ \mu_{\text{DI}}: \text{high rate of background deletions and inactivations, low rate of duplications} \]

\[ \mu_{\text{TRev}}: \text{background mutations as } \mu_{\text{DI}} + \text{evolved proportion of transcriptional mutations} \]
Slightly larger rate of duplications compensate for deletions and inactivations. Removing rRNA genes minimally affect growth.

Analogous to yeast.
Mutation rates evolve: Variable environment

Random switches $S_{in}=10 - 50$ (avrg. 250 generations)

$\mu_{BG}$: Only background mut.  
(dupl. del. inact. = 0.00333)

$\mu_{TR}$: Transcr. induced mutations  
(background mut. as in $\mu_{BG}$).

Evolved rates:  
Duplication = 0.0491  
Deletion = 0.0495  
Inactivation = 0.0003
Short vs. long term fitness:
- Mutations that are beneficial in the short run → accumulation mutational load in the long run
- Evolution reduces mutational load by means of larger rate of biased transcription-induced mutations

Long term evolution avoids evolutionary deterioration of genome (evo-evo)

Future work: Generality? Bacteria?
Acknowledgements:

Paulien
Renske
Jan-Kees

Members of Evo-evo collaboration