Nearly-neutral dynamics: when is selection not selective?

Daniel Lawson
Bioinformatics and Statistics Scotland
Macaulay Institute, Aberdeen

work with:

Henrik Jeldtoft Jensen
Imperial College London
Question: is evolution really stochastic?

• Focus on Bacteria

• Large population sizes (>10^{10})

• Weak selection acts on whole genome
  • Stability of protein folding (Aita, Ota and Usimi, 2003, JTB 221:599-613)
  • Gene regulation (Ohta, 2002, PNAS 99:5916134-16137)
  • Non coding DNA! (Ohta, 1997, Gene 205:261-267)
  • Genome length, reading fidelity, etc…

• How small should selection be for this not to matter?
Standard approach

• Consider allele A and mutant type a
  • Fitness of A is 1. Fitness of a is 1-s
  • Population size N
  • Initial population of a is 1

• Gamblers ruin:
  • Probability of winning relative to neutral case
    \[ \frac{p(s)}{p(0)} = 1 - Ns + O([Ns]^2) \]

• Therefore \( s^* \sim \frac{1}{N} \)
Standard approach

- Selection $O(10^{-10})$ is important!
- Theory: Bacteria shouldn’t evolve neutrally
- Genetics data: neutral evolution occurs
  - Neutral at some loci
  - Neutral phylogenetic tree
- Also: genetic inference relies on this!
- Theory or data interpretation is wrong
  - Inappropriate to apply theory?
  - Stochastic interpretation of evolution?
- Bacteria perhaps are more selected?
  - $\sim50\%$ of genes are selected in bacteria (Charlesworth and Eyre-Walker, 2006, Mol. Biol. Evol. 23:1348-1356)
  - 0-30% in apes
Are bacteria special?

• Form colonies: lower effective population

• High mutation rates: high diversity

• (mostly) clonal: gene interactions important
Fitness and reproduction

- "Fitness" concept
  - Doesn’t exist!

- Fitness assigned to genes?
  - Only when recombination is high

- Fitness assigned to whole types?
  - Better when recombination is low

- Difference between asexual and sexual reproduction
High mutation rates

• Create a distribution of many types
• Types don’t have to be very good to mutate again
• Gene interactions:
  • Impact of mutation depends on genotype
  • Multiple bad mutations can produce fit type
• How does this distribution effect evolution?
Evolution model

• N individuals reproduce clonally in a type space
  • Chosen randomly for death
  • Chosen proportionally to fitness for reproduction
  • Mutation: change at a single locus

• Mutation: maximum change in fitness = s
Fitness Landscapes

- Random
- Linear
- Top hat
Results

Width: standard deviation of genetic distribution

$S^* \sim N^{-\alpha}$
Results
Summary of Results

- Neutral region size:
  - Linear landscape: large region ($\alpha = 1.5$)
  - Top hat: as low mutation case ($\alpha = 1$)
  - Random fitness:
    - $\alpha \to 1$ as $p_{mutation} \to 0$ or dimension $D \to 0$
    - $\alpha \to 0$ as dimension $D \to \infty$
    - When $\alpha = 0$ an infinitely large population can still be neutral!
Why neutral evolution?

N=100, s*=0.01
Real evolution

• Selection “weaker” in large populations if:
  • Gene interaction is strong
  • Mutation rates per generation high

• Can form “Neutral Network”
  • Not all mutations are neutral
  • But population overcomes bad mutations by chance

• MAY apply to sexual population if gene interactions strong enough (NK landscape)
Caveats

• “Fitness Landscapes” don’t really exist
• Model not explicitly for DNA
• Only mutations with weak effect count as “dimensions”
Conclusion

• Combination of
  • High mutation rates
  • Gene interactions important over evolutionary time
  • Large space of possible mutations

• Allows statistically neutral evolution even when selection is present

• Occurring on neutral networks

• This happens all the time!
Implications

• Recombining/non-recombining regions and phenotypes/genotypes evolve differently
• Stochastic evolution in large populations is reasonable
  • Current genetic inference techniques supported!
• Selection at single mutation DOES NOT imply selection in population!
Further reading

This work:
“The role of weak selection and high mutation rates in nearly neutral evolution”
Daniel John Lawson and Henrik Jeldtoft Jensen (to appear)

Cited Literature:
Aita, Ota and Uusimi, 2003, JTB 221:599-613
Ohta, 2002, PNAS 99:5916134-16137
Ohta, 1997, Gene 205:261-267
Nearly neutral tree