MCMC INTEGRATION OVER EVOLUTIONARY HISTORIES OF METABOLIC NETWORKS

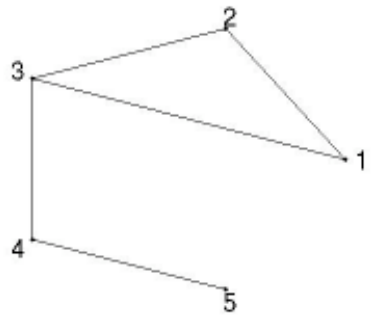
Hugo Fernandes
Hugo Martins
Ricardo Pinho
Rodrigo Liberal

IGC, 2008

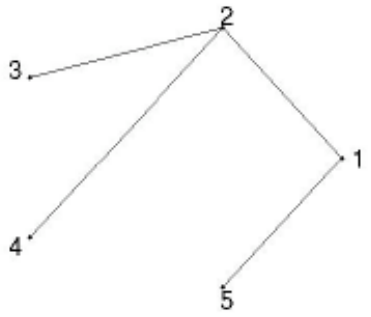
The Project

- Goals
  - Phylogenetic trees aka Probability of data $P_t(N1,N2)$

N1



N2

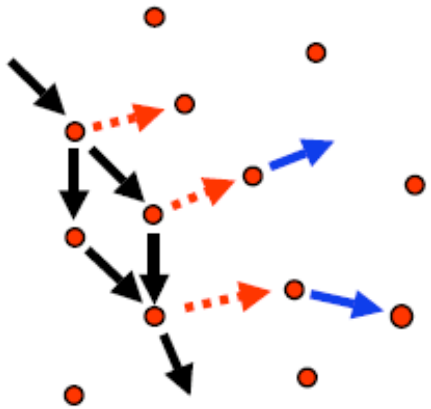


- Motivation
  - sequence alignment vs metabolic network alignment

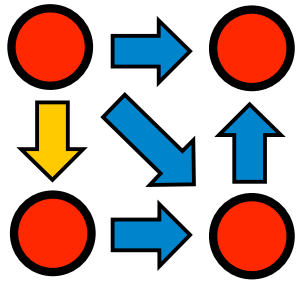
    (via evolutionary models)

Metabolic Pathways as graphs

nodes = metabolites, edges = reactions

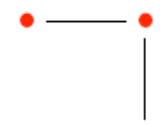


Black arrows:
'core'
metabolism
Red arrows:
'deletable' edges
Blue arrows:
'addable' edges



Markov chains

- Evolution of Metabolic Pathways as a Markovian process $P_t(N_1,N_2) = P(N_1)P_t(N_1 \rightarrow N_2)$

Network	Graph Structure
N_1	
N_2	
N_3	
N_4	

$$\frac{dP(t)}{dt} = QP(t) \qquad P(t) = e^{Qt}$$

$Q(N_i \rightarrow N_j)$	N_1	N_2	N_3	N_4	...
N_1	$-\sum_j q_{1j}$	λ	λ	0	...
N_2	μ	$-\sum_j q_{2j}$	0	λ	...
N_3	μ	0	$-\sum_j q_{3j}$	λ	...
N_4	0	μ	μ	$-\sum_j q_{4j}$	...
...	...	...	...	...	...

Computationally unfeasible

n	Number of all graphs with n nodes	Number of states
1	1	1
2	2	2
3	8	8
4	64	61
5	1024	969
6	32768	31738
7	2097152	2069964
8	268435456	267270033
9	68719476736	68629753641
10	35184372088832	35171000942698

Table 1: Number of possible states of the Markov Chain for networks of size n.

0. Network Parsimony

MCMCN!

- Assessing distances between networks
- (Number of indels, time derived from insertion and deletion rates)

→ (white board)

1. Monte Carlo

- Monte Carlo methods to deal with large metabolic networks

MCMCMHMB!

→ (white board)

2. Sampling sub-graphs

MCMCWN!

→ (white board)

Future Prospects

- Weighting the edges that stray from the core metabolism
- Fitness - better alignment or evolution for evolution's sake (without alignment)
- Parameter estimation
- Non uniform distributions (e.g. preferential attachment)
- Chemical equilibrium constant
- Meta-evolubility: Meta graph with meta grammar