

Generation of Gene Family Histories

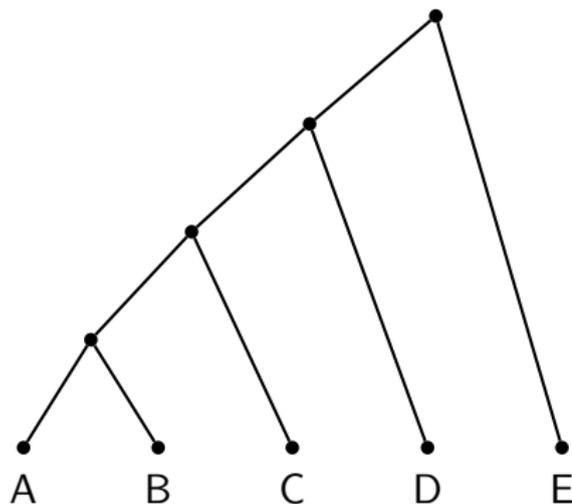
Maribel Hernandez Rosales
maribel@bioinf.uni-leipzig.de

February 16, 2011

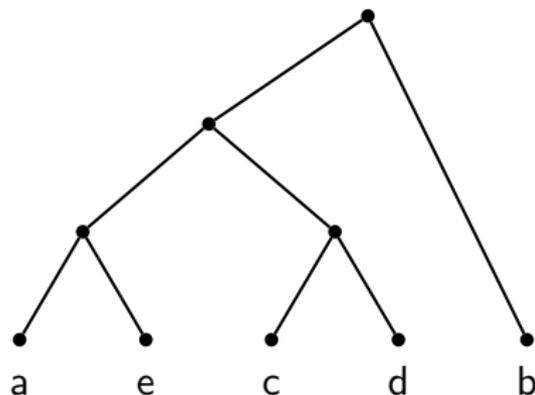
- 1 Background
 - Inconsistent Trees
 - Reconciliation
- 2 The Generation Method
 - Simulating Gene Family Histories
- 3 Reconciliation Tree
 - The Reconciled Tree
 - Orthologous Genes
- 4 Orthology & Homology
 - Orthology: Wake up! That was just a dream!
 - Generating Homologies
- 5 Test your Reconstruction!

Species Tree vs. Gene Tree

Species Tree



Gene Tree



Goodman et al. (1979) *Systematic Zoology*

Reconciliation

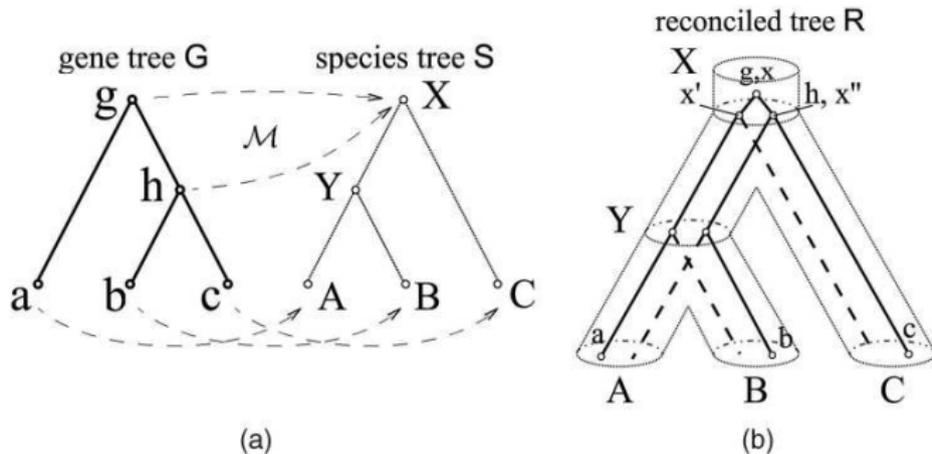
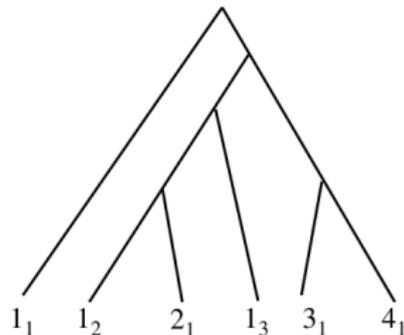


Figure: (a) Gene Tree G and Species Tree S are comparable by the leaf-mapping. (b) R is the reconciled tree for G and S

Bansal et al. (2009) EEE/ACM Transactions on Computational Biology and

Disadvantages of existing Reconciliation Methods

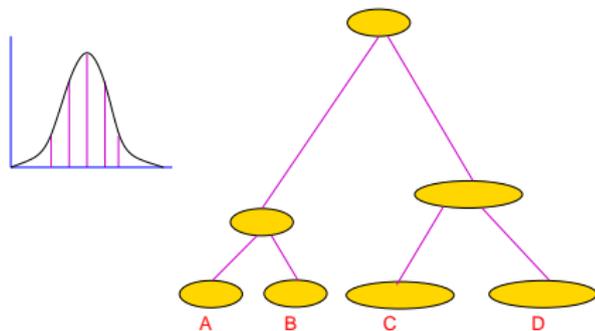
- 1 Only one gene per species in the gene tree. For more than one gene per species, many gene trees.
- 2 Whenever a gene duplication is placed a gene loss must be placed as well.
- 3 When in one gene tree there are more than 1 gene per species, the gene tree becomes huge and inaccurate.



A gene tree with more than one gene per species
(Chauve et al. 2007)

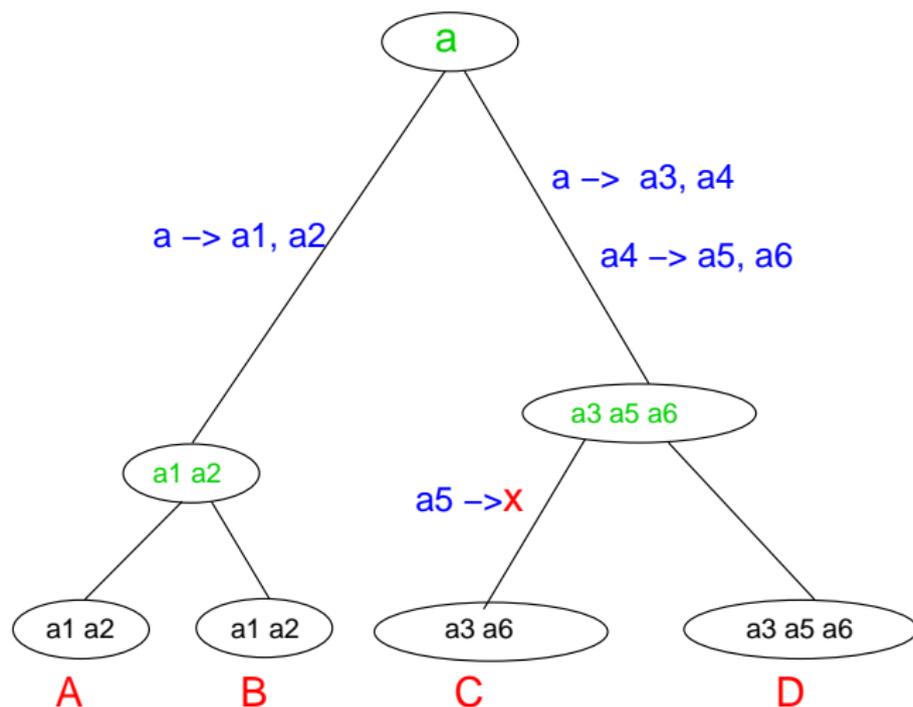
Species Tree

- 1 Species tree generated according to the Age Model
- 2 Length of edges lie in a normal distribution (parameter)

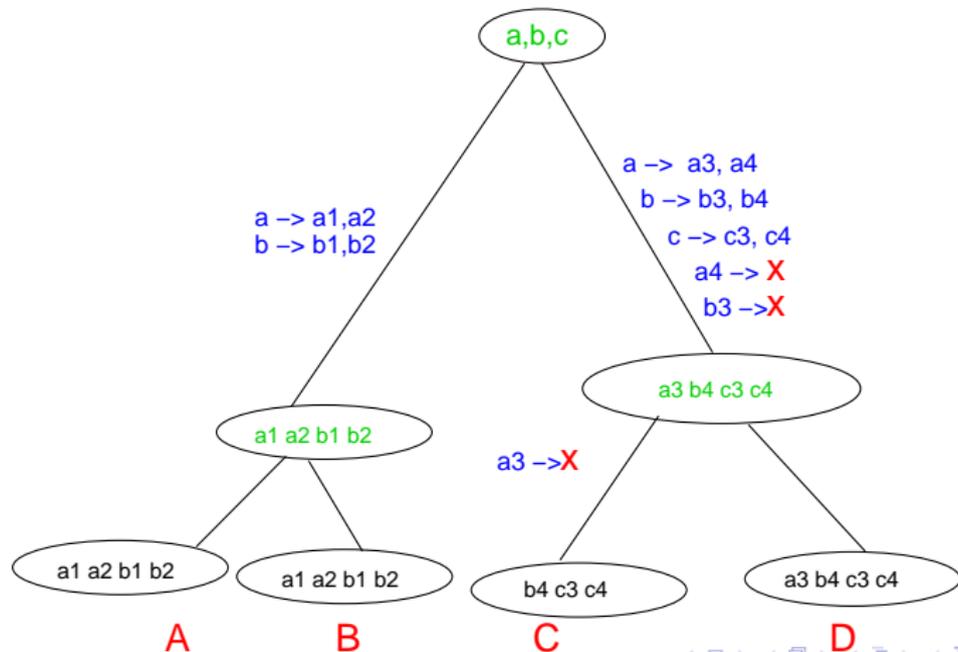


- 3 Generations of events according to a Poisson Process
- 4 Allowed events:
 - 1 Gene Duplication
 - 2 Cluster Duplication
 - 3 Genome Duplication
 - 4 Gene Loss

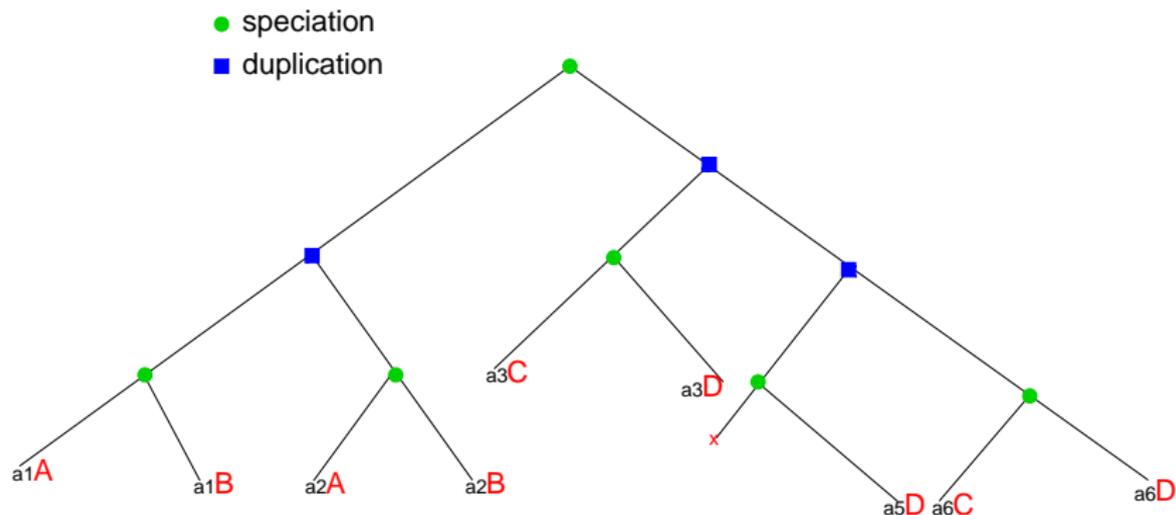
A one gene family history



Cluster Duplication, Genome Duplication, Forcing Gene Losses

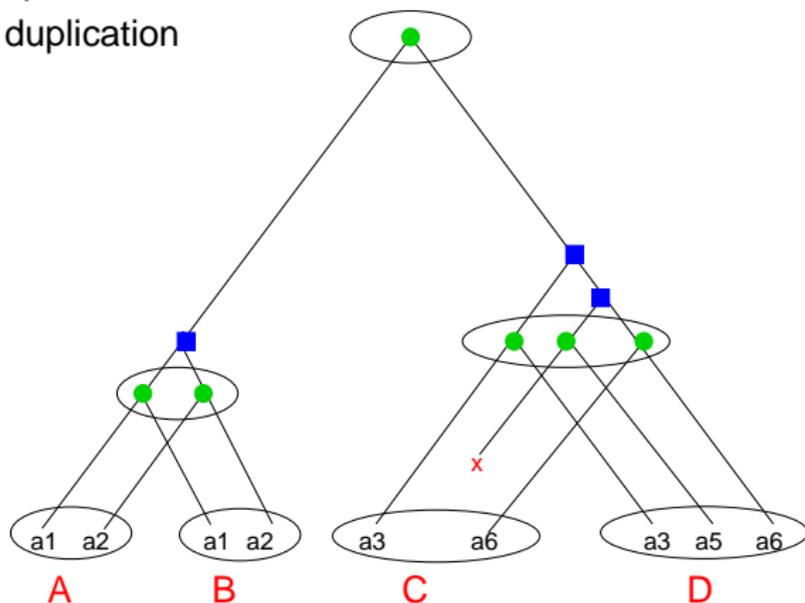


The Gene Tree



The Reconciled Tree

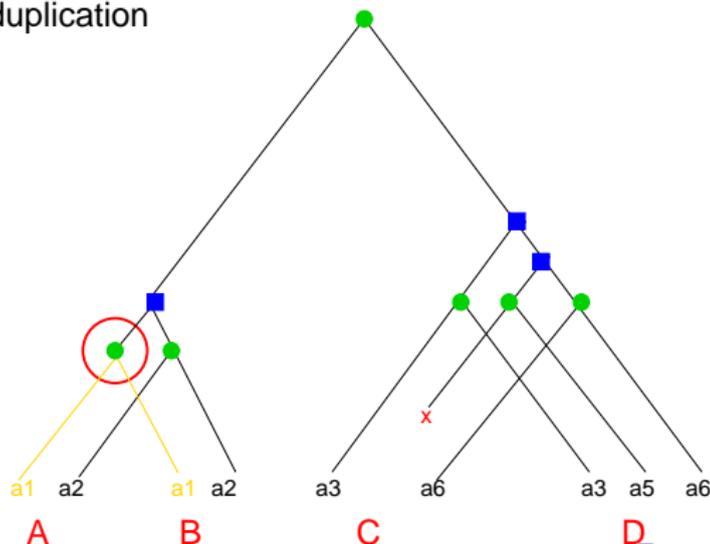
- speciation
- duplication



Orthologous Genes

Definition: Two genes are orthologs if their Lowest Common Ancestor in the Reconciled Tree is a Speciation event.

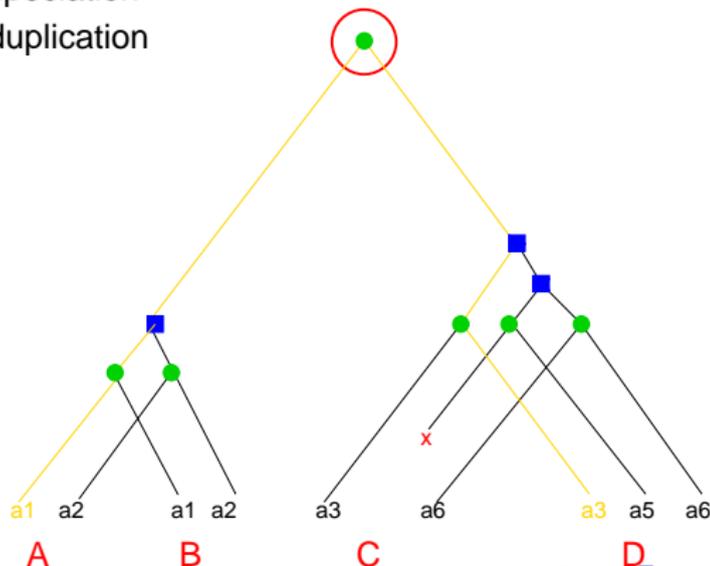
- speciation
- duplication



Orthologous Genes

Definition: Two genes are orthologs if their Lowest Common Ancestor in the Reconciled Tree is a Speciation event.

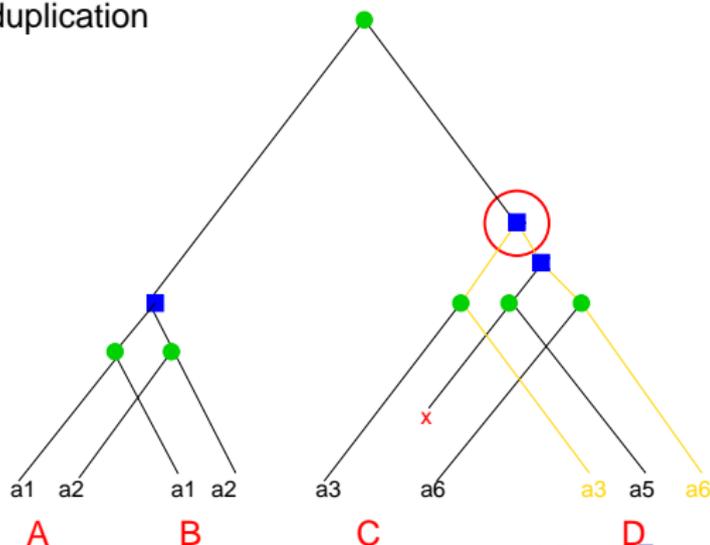
- speciation
- duplication



Orthologous Genes

Definition: Two genes are NOT orthologs if their Lowest Common Ancestor in the Reconciled Tree is a Duplication event.

- speciation
- duplication

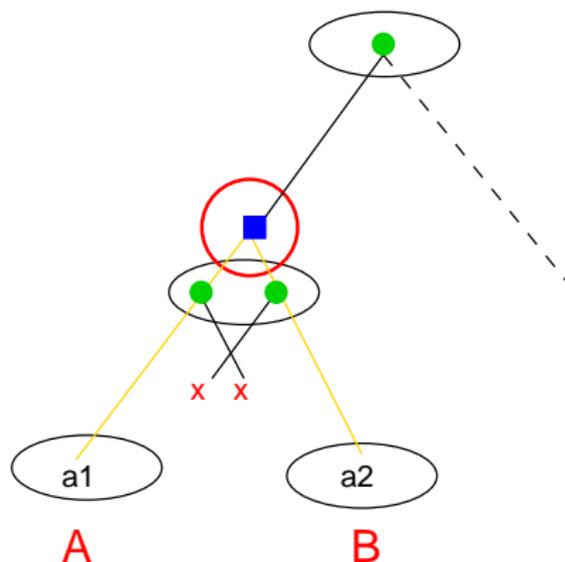


The Matrix of Orthology

	a1A	a2A	a1B	a2B	a3C	a6C	a3D	a5D	a6D
a1A	1	0	1	0	1	1	1	1	1
a2A	0	1	0	1	1	1	1	1	1
a1B	1	0	1	0	1	1	1	1	1
a2B	0	1	0	1	1	1	1	1	1
a3C	1	1	1	1	1	0	1	0	0
a6C	1	1	1	1	0	1	0	0	1
a3D	1	1	1	1	1	0	1	0	0
a5D	1	1	1	1	0	0	0	1	0
a6D	1	1	1	1	0	1	0	0	1

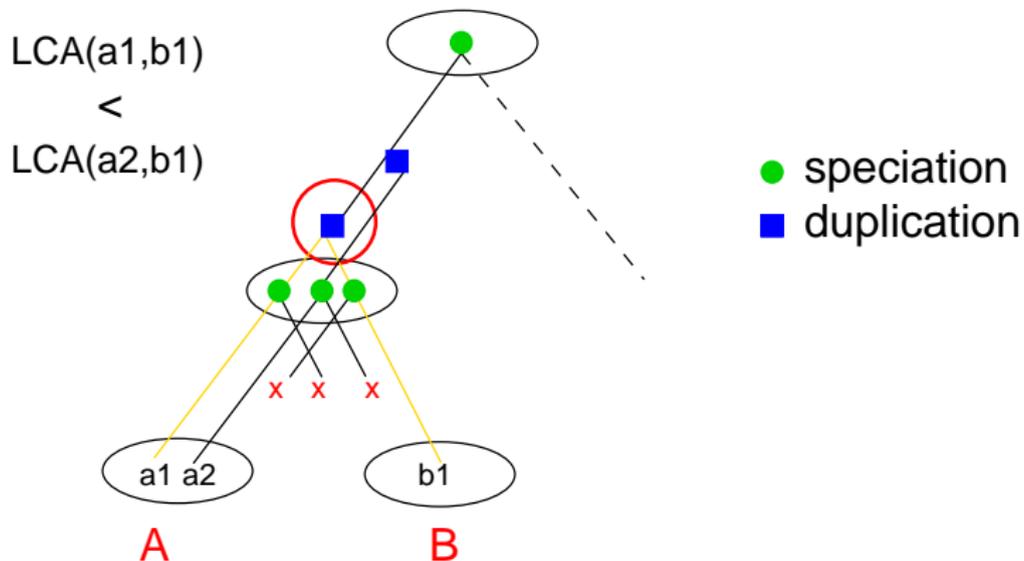
Real Life: Homology

Definition: For every gene in species A there must be a homologous gene in another species B



- speciation
- duplication

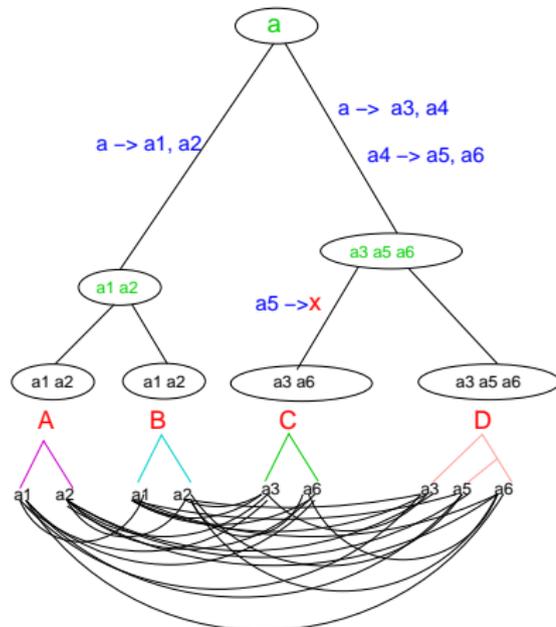
Choosing the closest related homologous gene



Test your Reconstruction!

We provide you...

- Gene Trees: one gene tree per species and/or
- One general gene tree with all genes for all species
- Matrix of homology
- Matrix of orthology
- The Reconciled Tree



Thanks to Nic, Marc, Markus and Peter for helpful discussions.

Thank you all for your attention!