Orthology Analysis
part of “Graphen und Netzwerke in der Biologie”

Sonja Prohaska

Computational EvoDevo
University Leipzig

Leipzig, SS 2011
“A is homologous to B” means...

- A and B are characters (nucleotide sequences, organs,... ) of individual organisms \( I_A \) and \( I_B \), respectively
- A and B are derived from a (last) common ancestor \( lca(A, B) \) by descent
- if A is a homolog of B, B is a homolog of A (symmetric relation)
- “homologous” in respect to syntax (structure) or semantics (function)?
- does NOT just mean “A is similar to B”
- similarity might hint at homology
- similarity without common ancestry is called analogy
if $A$ and $B$ driven from the $lca(A, B)$ by **duplication**, $A$ and $B$ are **paralogous**

if $A$ and $B$ are homologs and $l_A = l_B$ than they are **in-paralogs**

if $A$ is paralogous to $B$ and $l_A \neq l_B$ than they are **out-paralogs**

if $A$ and $B$ driven from the $lca(A, B)$ by **speciation**, $A$ and $B$ are **orthologous**
Orthology – Paralogy

Sonja Prohaska

Orthology Analysis
1:1, 1:many, many:many orthology

1:1 orthology

many:1 orthology

many:many orthology
More information from additional species and about timing of gene duplication and speciation events can change the view.

\[
\begin{align*}
d((a:X),(b:X)) &= 2t \\
d((a:X),(c:X)) &= 2t \\
d((b:X),(c:X)) &= 2s \\
d((a:X2'),(c:X2'')) &= 2t \\
d((b:X1),(c:X2'')) &= 2u \\
d((a:X2'),(b:X1)) &= 2u
\end{align*}
\]