Phylogenetic incongruence through the lens of Monadic Second Order logic

Steven Kelk, Leo van Iersel, Celine Scornavacca, Mathias Weller

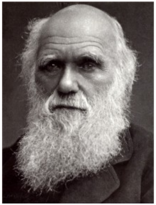
ISE-M, Equipe Phylogénie & Evolution Moléculaires

GROW 2015

From Aristotle to Darwin

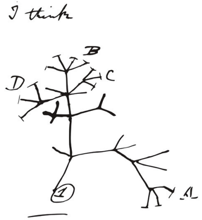
Since Aristotle, naturalists have always tried to classify the abundance of creatures that populate the Earth.

- Aristote: the *scala naturae*;
- Carl von Linné: classification of living;
- Antoine Laurent de Jussieu;
- Leclerc de Buffon: the first to evoke the possibility that species can evolve;
- Jean-Baptiste Lamarck: first theory of evolution;
- Charles Darwin: *The Origins of Species* (1859).



From ‘The Origin of Species”

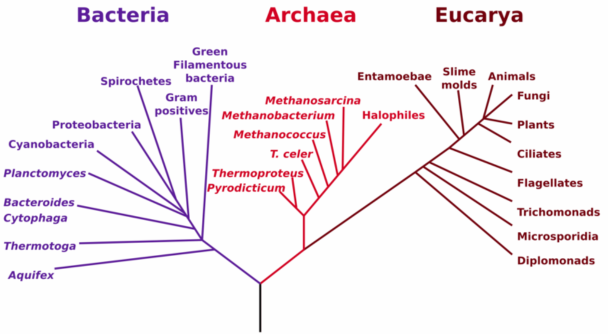
- It is a truly wonderful fact ... that all animals and all plants throughout all time and space should be related to each other in groups, subordinate to groups. [...]
- The affinities of all the beings of the same class have sometimes been represented by a great tree. [...] The green and budding twigs may represent existing species; and those produced during former years may represent the long succession of extinct species.



Charles Darwin, (1872), pp. 170-171. The Origin of Species. Sixth Edition. The Modern Library, New York.

Phylogenetics

Phylogenetics aims at clarifying, using molecular and morphological data, the evolutionary relationships that exist among different species. These relationships can be represented through phylogenetic trees or phylogenies (AIM: the TOL – Tree Of Life).

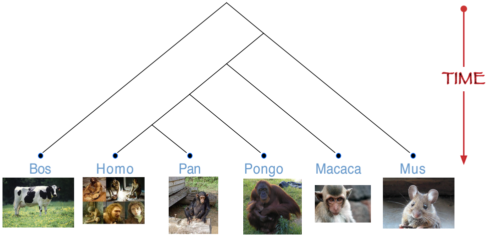


Woese 1987; Barns et al. 1996; Brown et Doolittle 1997

Rooted phylogenetic trees ...

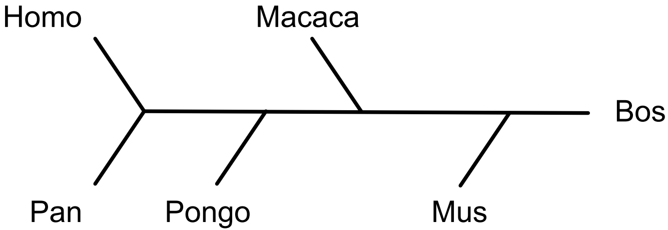
... are out-branching trees with no indegree-1 outdegreee-1 nodes, where sinks are associated to a set of species:

- the sinks or taxa represent existing organisms
- the only node with indegree-0 is called root
- internal nodes represent hypothetical ancestors
- each internal node represents the lowest common ancestor of all taxa below it (clade)
- nodes and branches can have several kinds of information associated with them, such as time or amount of evolution estimates.



Unrooted phylogenetic trees ...

... are trees with no degree-2 nodes, where leaves are associated to a set of species.



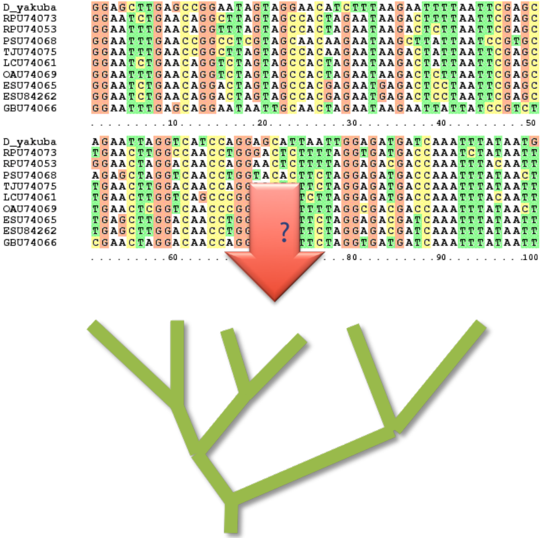
Phylogenetics reconstruction

With the discovery of DNA by Watson and Crick in 1953 and the design of sequencing techniques, a new kind of information became available: molecular data.

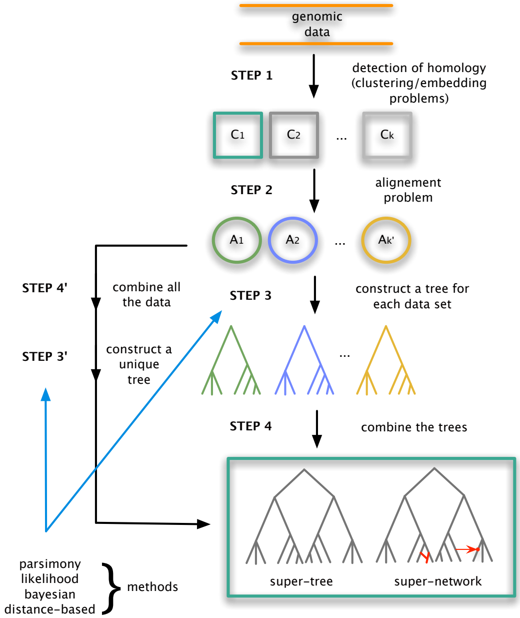
Today, phylogenies are obtained by studying:

- discrete characters;
- molecular sequences;
- gene frequencies;
- restriction sites;
- microsatellites;
- ...

Molecular phylogenetics

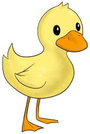


The 4 big steps of phylogenetics reconstruction



Gene trees

Gene trees are built by analyzing a gene family, i.e., homologous molecular sequences appearing in the genome of different organisms

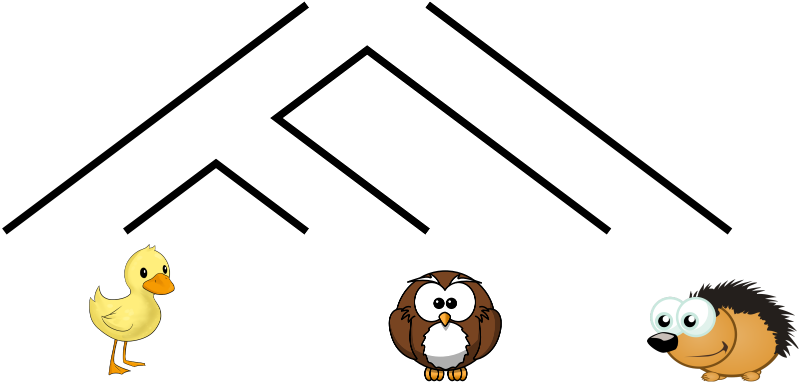


```
ACGTGCTTCGTCACCGTGACTGATCGTGCTAGCT
CTGTGACTGATCGTCTGATCGATGCATCATCTAA
```

```
TGCACCGACGATTGGATTGCTGTCAGCCTACGA
TTATTCTCGATGTTCCTTCTGACCGATGCTGAC
```

```
CGATCACTTAGAGCTGAGCTGGCGTCGTGAGCCT
AGCTGCGCGCGTGCTGATCCTAGCTAGTCGCTGA
```

Gene trees

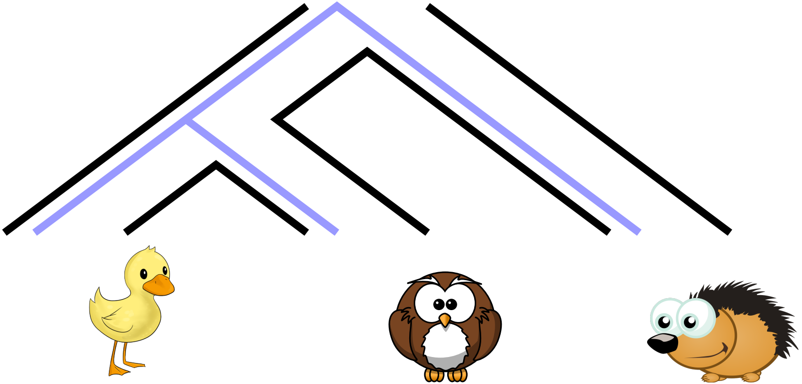


ACGTGCTTCGTCACCGTGACTGATCGTGCTAGCT
CTGTGACTGATCGTCTGATCGATGCATCATCTAA

TGCACCGACGATTGGATTGCTGTCAGCCTACGA
TTATTCTCGATGTTCCTTCTGACCGATGCTGAC

CGATCACTTAGAGCTGAGCTGGCGTCGTGAGCCT
AGCTGCGCGCGTGCTGATCCTAGCTAGTCGCTGA

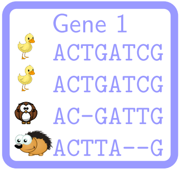
Gene trees



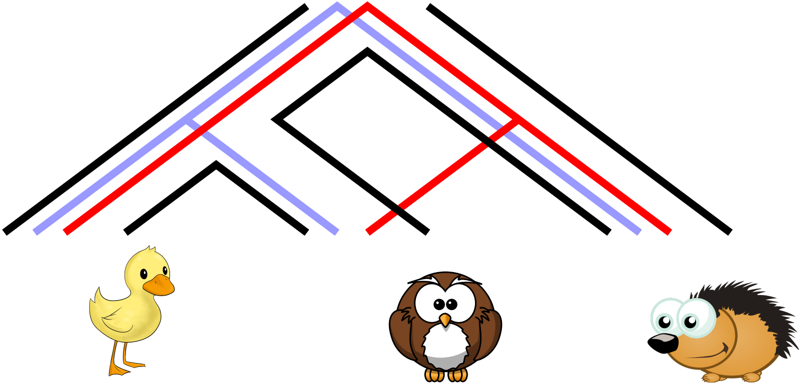
ACGTGCTTCGTCACCGTGACTGATCGTGCTAGCT
CTGTGACTGATCGTCTGATCGATGCATCATCTAA

TGCACCGACGATTGGATTGCTGTCAGCCTACGA
TTATTCTCGATGTTCCTTCTGACCGATGCTGAC

CGATCACTTAGAGCTGAGCTGGCGTCGTGAGCCT
AGCTGCGCGCGTGCTGATCCTAGCTAGTCGCTGA



Gene trees



ACGTGCTTCGTCACCGTGACTGATCGTGCTAGCT
CTGTGACTGATCGTCTGATCGATGCATCATCTAA

TGCACCGACGATTGGATTGCTGTCAGCCTACGA
TTATTCTCGATGTTCCTTCTGACCGATGCTGAC

CGATCACTTAGAGCTGAGCTGGCGTCGTGAGCCT
AGCTGCGCGCGTGCTGATCCTAGCTAGTCGCTGA

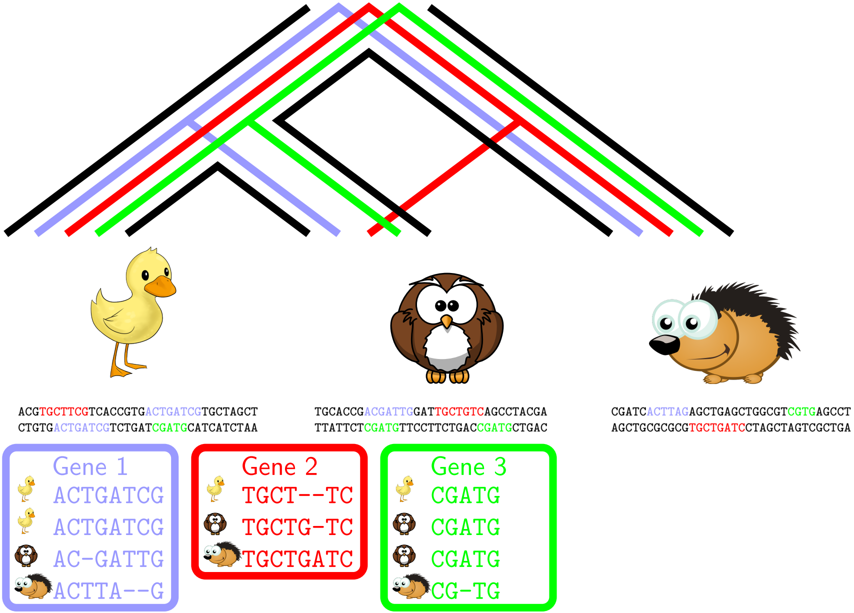
Gene 1

ACTGATCG
ACTGATCG
AC-GATTG
ACTTA--G

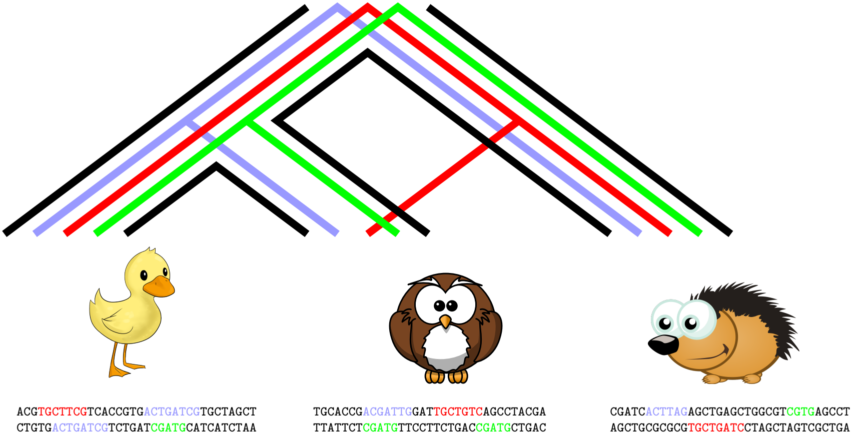
Gene 2

TGCT--TC
TGCTG-TC
TGCTGATC

Gene trees


ACGTGCTTCGTCACCGTGACTGATCGTGCTAGCT
CTGTGACTGATCGTCTGATCGATGCATCATCTAA
TGCACCGACGATTGGATTGCTGTCAGCCTACGA
TTATTCTCGATGTTCCTTCTGACCGATGCTGAC
CGATCACTTAGAGCTGAGCTGGCGTCGTGAGCCT
AGCTGCGCGCGTGCTGATCCTAGCTAGTCGCTGA
Gene 1
ACTGATCG
ACTGATCG
AC-GATTG
ACTTA--G
Gene 2
TGCT--TC
TGCTG-TC
TGCTGATC
Gene 3
CGATG
CGATG
CGATG
CG-TG

Gene trees



Gene trees can significantly differ from the species tree for:

- methodological reasons
- biological reasons

How to compare/combine them?

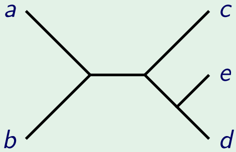
Agreement Forests

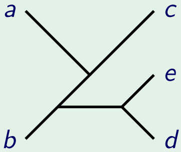
Agreement forest = subforest of T_1 and T_2 (which they agree on)

Agreement Forests

Agreement forest = subforest of T_1 and T_2 (which they agree on)

Unrooted Agreement Forests

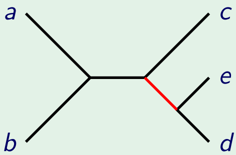


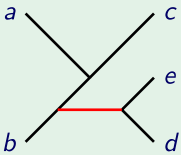


Agreement Forests

Agreement forest = subforest of T_1 and T_2 (which they agree on)

Unrooted Agreement Forests

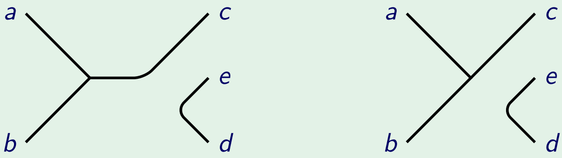




Agreement Forests

Agreement forest = subforest of T_1 and T_2 (which they agree on)

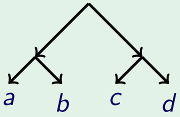
Unrooted Agreement Forests

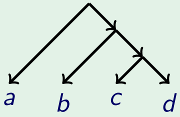


Agreement Forests

Agreement forest = subforest of T_1 and T_2 (which they agree on)

Rooted Agreement Forests

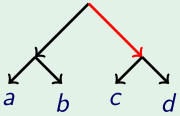


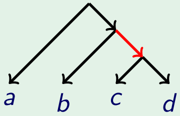


Agreement Forests

Agreement forest = subforest of T_1 and T_2 (which they agree on)

Rooted Agreement Forests

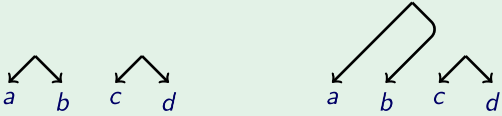




Agreement Forests

Agreement forest = subforest of T_1 and T_2 (which they agree on)

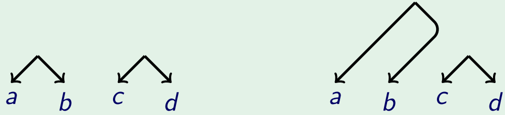
Rooted Agreement Forests



Agreement Forests

Agreement forest = subforest of T_1 and T_2 (which they agree on)
Maximum (Un)rooted Agreement Forest (uMAF/rMAF): #components $\rightarrow$ min

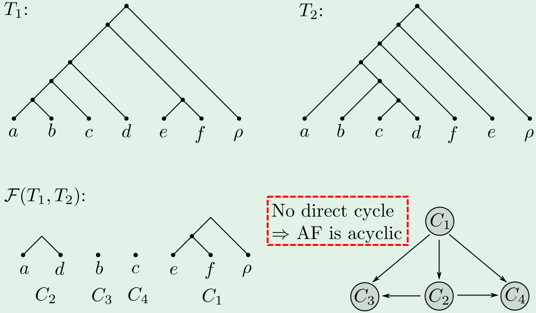
Rooted Agreement Forests



Agreement Forests

Agreement forest = subforest of T_1 and T_2 (which they agree on)
Maximum (Un)rooted Agreement Forest (uMAF/rMAF): #components → min

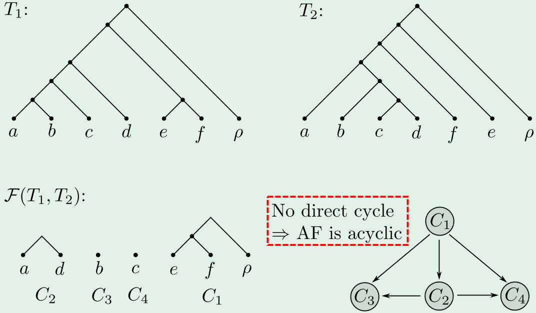
Acyclic Agreement Forests



Agreement Forests

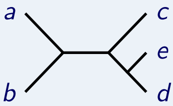
Agreement forest = subforest of T_1 and T_2 (which they agree on)
Maximum (Un)rooted Agreement Forest (uMAF/rMAF): #components $\rightarrow$ min
Maximum Acyclic Agreement Forest (MAAF): acyclic components

Acyclic Agreement Forests



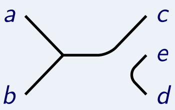
Phylogenetic Tree Distances

Tree Bisection & Reconnect



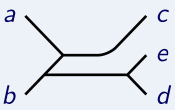
Phylogenetic Tree Distances

Tree Bisection & Reconnect



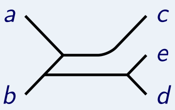
Phylogenetic Tree Distances

Tree Bisection & Reconnect



Phylogenetic Tree Distances

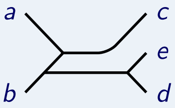
Tree Bisection & Reconnect



TBR distance:
min #TBR moves

Phylogenetic Tree Distances

Tree Bisection & Reconnect



TBR distance:
min #TBR moves

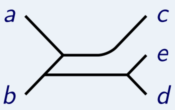
TBR & Agreement Forests

TBR dist=$|\text{uMAF}| - 1$

[Allen & Steel, 2001]

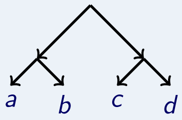
Phylogenetic Tree Distances

Tree Bisection & Reconnect



TBR distance:
min #TBR moves

rt'ed Subtree Prune & Regraft



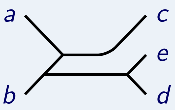
TBR & Agreement Forests

TBR dist$=|$uMAF$| - 1$

[Allen & Steel, 2001]

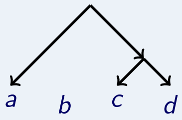
Phylogenetic Tree Distances

Tree Bisection & Reconnect



TBR distance:
min #TBR moves

rt'ed Subtree Prune & Regraft



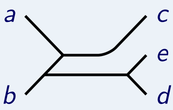
TBR & Agreement Forests

TBR dist=$|\text{uMAF}| - 1$

[Allen & Steel, 2001]

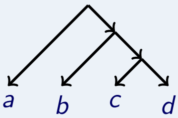
Phylogenetic Tree Distances

Tree Bisection & Reconnect



TBR distance:
min #TBR moves

rt'ed Subtree Prune & Regraft



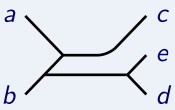
TBR & Agreement Forests

TBR dist$=|$uMAF$| - 1$

[Allen & Steel, 2001]

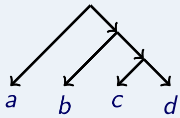
Phylogenetic Tree Distances

Tree Bisection & Reconnect



TBR distance:
min #TBR moves

rt'ed Subtree Prune & Regraft



rSPR distance:
min #rSPR moves

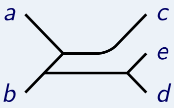
TBR & Agreement Forests

TBR dist=$|$uMAF$| - 1$

[Allen & Steel, 2001]

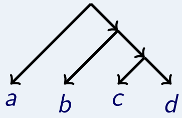
Phylogenetic Tree Distances

Tree Bisection & Reconnect



TBR distance:
min #TBR moves

rt'ed Subtree Prune & Regraft



rSPR distance:
min #rSPR moves

TBR & Agreement Forests

TBR dist$=|$uMAF$|-1$

[Allen & Steel, 2001]

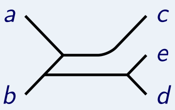
rSPR & Agreement Forests

rSPR dist$=|$rMAF$|-1$

[Bordewich & Semple, 2004]

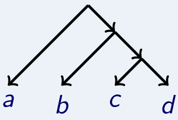
Phylogenetic Tree Distances

Tree Bisection & Reconnect



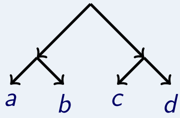
TBR distance:
min #TBR moves

rt'ed Subtree Prune & Regraft



rSPR distance:
min #rSPR moves

Hybridization Number



TBR & Agreement Forests

TBR dist=|uMAF| − 1

[Allen & Steel, 2001]

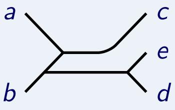
rSPR & Agreement Forests

rSPR dist=|rMAF| − 1

[Bordewich & Semple, 2004]

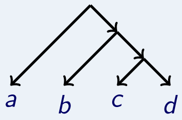
Phylogenetic Tree Distances

Tree Bisection & Reconnect



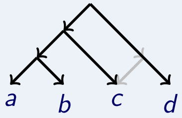
TBR distance:
min #TBR moves

rt'ed Subtree Prune & Regraft



rSPR distance:
min #rSPR moves

Hybridization Number



TBR & Agreement Forests

TBR dist=|uMAF| − 1

[Allen & Steel, 2001]

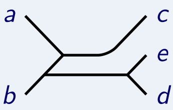
rSPR & Agreement Forests

rSPR dist=|rMAF| − 1

[Bordewich & Semple, 2004]

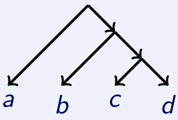
Phylogenetic Tree Distances

Tree Bisection & Reconnect



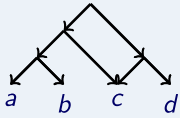
TBR distance:
min #TBR moves

rt'ed Subtree Prune & Regraft



rSPR distance:
min #rSPR moves

Hybridization Number



TBR & Agreement Forests

TBR dist=$|$uMAF$| - 1$

[Allen & Steel, 2001]

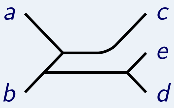
rSPR & Agreement Forests

rSPR dist=$|$rMAF$| - 1$

[Bordewich & Semple, 2004]

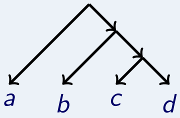
Phylogenetic Tree Distances

Tree Bisection & Reconnect



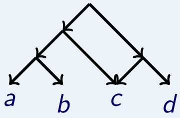
TBR distance:
min #TBR moves

rt'ed Subtree Prune & Regraft



rSPR distance:
min #rSPR moves

Hybridization Number



HN:
$\min_N$ #indeg-2 nodes

TBR & Agreement Forests

TBR dist$=|$uMAF$|-1$

[Allen & Steel, 2001]

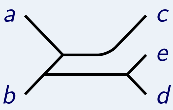
rSPR & Agreement Forests

rSPR dist$=|$rMAF$|-1$

[Bordewich & Semple, 2004]

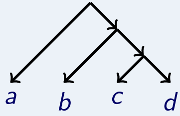
Phylogenetic Tree Distances

Tree Bisection & Reconnect



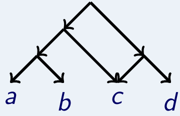
TBR distance:
min #TBR moves

rt'ed Subtree Prune & Regraft



rSPR distance:
min #rSPR moves

Hybridization Number



HN:
$\min_N \text{FES}\#(UG(N))$

TBR & Agreement Forests

TBR dist$=|\text{uMAF}| - 1$

[Allen & Steel, 2001]

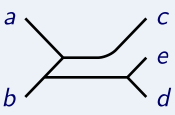
rSPR & Agreement Forests

rSPR dist$=|\text{rMAF}| - 1$

[Bordewich & Semple, 2004]

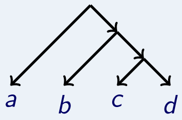
Phylogenetic Tree Distances

Tree Bisection & Reconnect



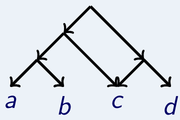
TBR distance:
min #TBR moves

rt'ed Subtree Prune & Regraft



rSPR distance:
min #rSPR moves

Hybridization Number



HN:
$\min_N \text{FES}\#(UG(N))$

TBR & Agreement Forests

TBR dist$=|$uMAF$| - 1$

[Allen & Steel, 2001]

rSPR & Agreement Forests

rSPR dist$=|$rMAF$| - 1$

[Bordewich & Semple, 2004]

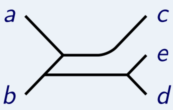
HN & Agreement Forests

HN$=|$MAAF$| - 1$

[Baroni et al., 2005]

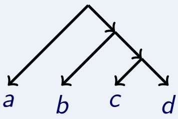
Phylogenetic Tree Distances

Tree Bisection & Reconnect



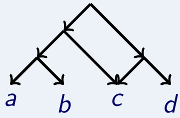
TBR distance:
min #TBR moves

rt'ed Subtree Prune & Regraft



rSPR distance:
min #rSPR moves

Hybridization Number



HN:
$\min_N \text{FES}\#(UG(N))$

TBR & Agreement Forests

TBR dist=|uMAF| − 1

[Allen & Steel, 2001]

rSPR & Agreement Forests

rSPR dist=|rMAF| − 1

[Bordewich & Semple, 2004]

HN & Agreement Forests

HN=|MAAF| − 1

[Baroni et al., 2005]

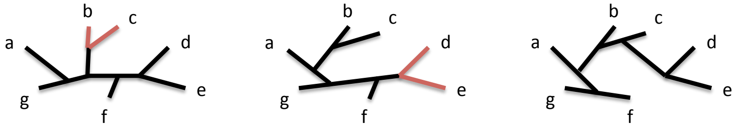
Complexity results

NP-hard [Allen & Steel, 2001, Bordewich & Semple, 2004 – 2007], but FPT in their natural parameterizations:

- $O(4^k \cdot n)$
- $O(2.42^k \cdot n)$ (they claim $O(2^k \cdot n)$ but paper not available yet)
- $O(3.18^k \cdot n)$ [Whidden, Beiko & Zeh, 2013]

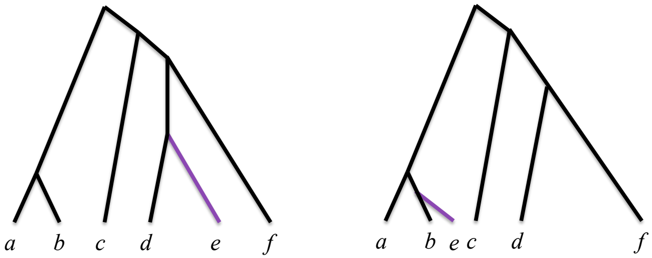
Biological motivation

- **TBR**: used to compare trees and studied to better understand how local-search heuristics, based on rearrangement operations, navigate the space of phylogenetic trees



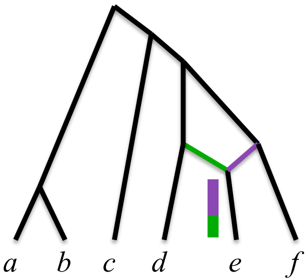
Biological motivation

- **TBR**: used to compare trees and studied to better understand how local-search heuristics, based on rearrangement operations, navigate the space of phylogenetic trees
- **rSPR**: the same as above, plus useful to count putative lateral gene transfers



Biological motivation

- **TBR**: used to compare trees and studied to better understand how local-search heuristics, based on rearrangement operations, navigate the space of phylogenetic trees
- **rSPR**: the same as above, plus useful to count putative lateral gene transfers
- **HN**: useful to count putative hybridization events



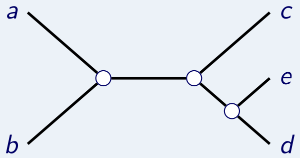
Computational motivation

tree pair	*taxa*	*HN*	*rSPR*	*TBR*
ndhF-phyB	40	14	12	6
ndhF-rbcL	36	13	10	6
ndhF-rpoC2	34	12	11	8
ndhF-waxy	19	9	7	4
ndhF-ITS	46	19	19	15
phyB-rbcL	21	4	4	4
phyB-rpoC2	21	7	6	4
phyB-waxy	14	3	3	2
phyB-ITS	30	8	8	7
rbcL-rpoC2	26	13	11	6
rbcL-waxy	12	7	6	3
rbcL-ITS	29	14	13	10
rpoC2-waxy	10	1	1	1
rpoC2-ITS	31	15	14	10
waxy-ITS	15	8	7	5

Table: Experiments on the *Poaceae* grass dataset

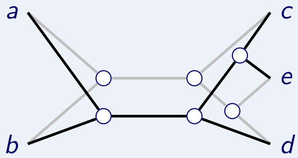
Display Graph

Unrooted



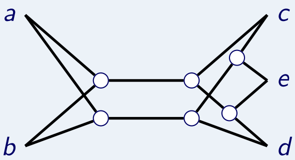
Display Graph

Unrooted



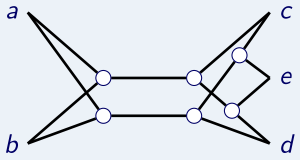
Display Graph

Unrooted

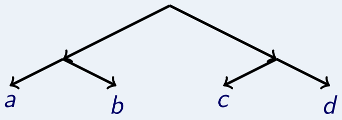


Display Graph

Unrooted

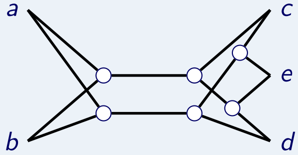


Rooted

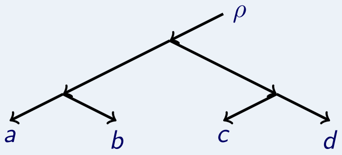


Display Graph

Unrooted

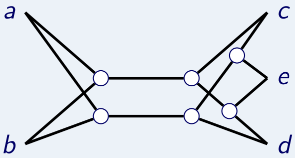


Rooted

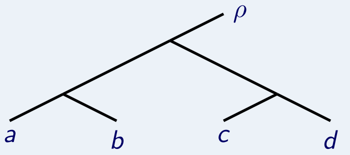


Display Graph

Unrooted



Rooted



MSOL Formulation

Theorem (Grigoriev, Kelk, Lekić, 2015)

The display graph of two agreeing trees has treewidth at most 2.

MSOL Formulation

Theorem (Grigoriev, Kelk, Lekić, 2015)

The display graph of two agreeing trees has treewidth at most 2.

$\rightsquigarrow$ tw(display graph) bounded in agreement forest sizes

MSOL Formulation

Theorem (Grigoriev, Kelk, Lekić, 2015)

The display graph of two agreeing trees has treewidth at most 2.

$\rightsquigarrow$ tw(display graph) bounded in agreement forest sizes $\rightsquigarrow$ Courcelle

MSOL Formulation

Theorem (Grigoriev, Kelk, Lekić, 2015)

The display graph of two agreeing trees has treewidth at most 2.

$\rightsquigarrow$ tw(display graph) bounded in agreement forest sizes $\rightsquigarrow$ Courcelle

uMAF ingredients

- root T_1 and T_2 arbitrarily
- represent edge deletion as their "lower" vertex
- leaves a, b in the same subtree w.r.t. solution K
  $\iff$ the a-b-path intersects K only in the LCA of a and b
- any 4 leaves in the same subtree induce the same topology in T_1 and T_2 $\rightsquigarrow$ agreement [Buneman, 1971]

MSOL Formulation

Theorem (Grigoriev, Kelk, Lekić, 2015)

The display graph of two agreeing trees has treewidth at most 2.

$\rightsquigarrow$ tw(display graph) bounded in agreement forest sizes $\rightsquigarrow$ Courcelle

rMAF ingredients

- represent edge deletion as their "lower" vertex
- leaves a, b in the same subtree w.r.t. solution K
  $\iff$ the a-b-path intersects K only in the LCA of a and b
- any 3 leaves in the same subtree induce the same topology in T_1 and T_2 $\rightsquigarrow$ agreement [Buneman, 1971]

MSOL Formulation

Theorem (Grigoriev, Kelk, Lekić, 2015)

The display graph of two agreeing trees has treewidth at most 2.

$\rightsquigarrow$ tw(display graph) bounded in agreement forest sizes $\rightsquigarrow$ Courcelle

MAAF ingredients

- represent edge deletion as their "lower" vertex
- leaves a, b in the same subtree w.r.t. solution K
  $\iff$ the a-b-path intersects K only in the LCA of a and b
- any 3 leaves in the same subtree induce the same topology in T_1 and T_2 $\rightsquigarrow$ agreement [Buneman, 1971]
- "corresponding"-relation linking the roots of the agreeing subtrees represented by K
- force acyclicity on this relation

MSOL Formulation

Theorem (Grigoriev, Kelk, Lekić, 2015)

The display graph of two agreeing trees has treewidth at most 2.

$\rightsquigarrow$ tw(display graph) bounded in agreement forest sizes $\rightsquigarrow$ Courcelle

MAF ingredients

- represent edge deletion as their "lower" vertex
- leaves a, b in the same subtree w.r.t. solution K
  $\iff$ the a-b-path intersects K only in the LCA of a and b
- any 3 leaves in the same subtree induce the same topology in T_1 and T_2 $\rightsquigarrow$ agreement [Buneman, 1971]
- "corresponding"-relation linking the roots of the agreeing subtrees represented by K
- force acyclicity on this relation

Theorem

Computing TBR-, rSPR-dist and HN is FPT in the treewidth of the display graph.

MSOL$_1$ Formulation

Theorem (Grigoriev, Kelk, Lekić, 2015)

The display graph of two agreeing trees has treewidth at most 2.

$\rightsquigarrow$ tw(display graph) bounded in agreement forest sizes $\rightsquigarrow$ Courcelle

MAF ingredients

- represent edge deletion as their "lower" vertex
- leaves a, b in the same subtree w.r.t. solution K
  $\iff$ the a-b-path intersects K only in the LCA of a and b
- any 3 leaves in the same subtree induce the same topology in T_1 and T_2 $\rightsquigarrow$ agreement [Buneman, 1971]
- "corresponding"-relation linking the roots of the agreeing subtrees represented by K
- force acyclicity on this relation

Theorem

Computing TBR-, rSPR-dist and HN is FPT in the cliquewidth of the display graph.

Experiments on the Poaceae grass dataset – 2

tree pair	*taxa*	*HN*	*rSPR*	*TBR*	**TW** $\leq$	*size display graph* vertices, edges
ndhF-phyB	40	14	12	6	**3**	118,156
ndhF-rbcL	36	13	10	6	**3**	106,140
ndhF-rpoC2	34	12	11	8	**5**	100,132
ndhF-waxy	19	9	7	4	**4**	55,72
ndhF-ITS	46	19	19	15	**6**	136,180
phyB-rbcL	21	4	4	4	**3**	61,80
phyB-rpoC2	21	7	6	4	**3**	61,80
phyB-waxy	14	3	3	2	**3**	40,52
phyB-ITS	30	8	8	7	**4**	88,116
rbcL-rpoC2	26	13	11	6	**5**	76,100
rbcL-waxy	12	7	6	3	**3**	34,44
rbcL-ITS	29	14	13	10	**5**	85,112
rpoC2-waxy	10	1	1	1	**3**	28,36
rpoC2-ITS	31	15	14	10	**6**	91,120
waxy-ITS	15	8	7	5	**4**	43,56

Table: Experiments on the *Poaceae* grass dataset. The "Greedy Fill-In" heuristic [Bodlaender & Koster, 2010] was used to compute an upper bound since exact computation of the treewidth was computationally infeasible.

Further Work

- Can we do better? $O(c^{tw})$, for a small constant?
- Can we find a “finer” bound on the treewidth w.r.t. the agreement forests size?
- Is it NP-hard to exactly compute *tw* on display graphs?
- Which patterns in the display graph (and thus in the trees) make the treewidth grows?
- Can we remove these patterns in the display graph and reduce its treewidth?
- ...
- ...

Now, a hint on our ongoing work for a practical algorithm

Ongoing Work: Towards a Practical Algorithm

Observation

$\exists$ *optimal tree decomposition with taxa in decomposition leaves*

Ongoing Work: Towards a Practical Algorithm

Observation

∃ *optimal tree decomposition with taxa in decomposition leaves*

Dynamic Programming Idea

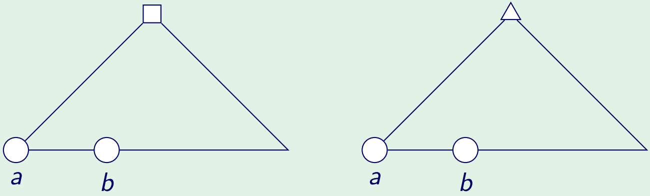


table: [X] = minimum #deletions "below" X

Ongoing Work: Towards a Practical Algorithm

Observation

$\exists$ *optimal tree decomposition with taxa in decomposition leaves*

Dynamic Programming Idea

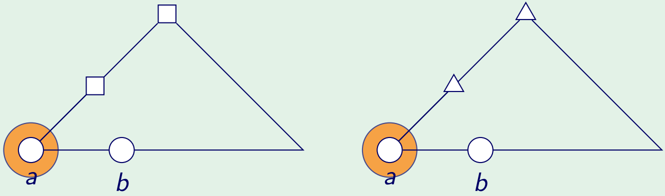


table: $[X \quad\quad]$ = minimum #deletions "below" X

Ongoing Work: Towards a Practical Algorithm

Observation

∃ *optimal tree decomposition with taxa in decomposition leaves*

Dynamic Programming Idea

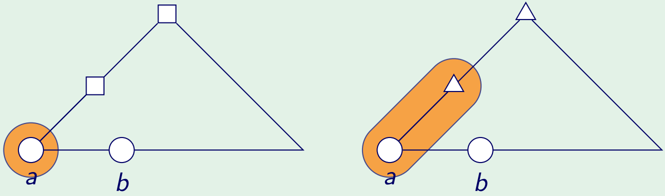


table: [X] = minimum #deletions "below" X

Ongoing Work: Towards a Practical Algorithm

Observation

∃ *optimal tree decomposition with taxa in decomposition leaves*

Dynamic Programming Idea

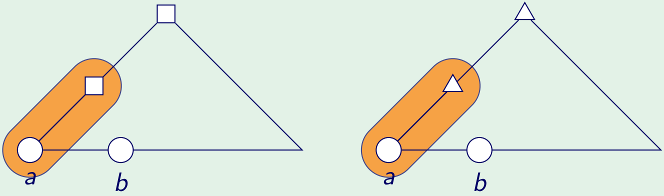


table: $[X \quad\quad] =$ minimum #deletions "below" X

Ongoing Work: Towards a Practical Algorithm

Observation

$\exists$ *optimal tree decomposition with taxa in decomposition leaves*

Dynamic Programming Idea

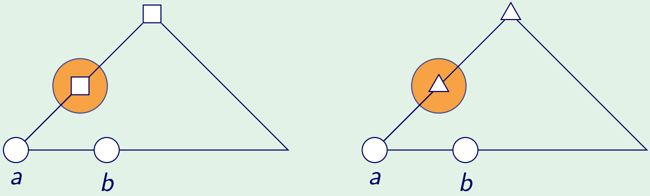


table: $[X \qquad] =$ minimum #deletions "below" X

Ongoing Work: Towards a Practical Algorithm

Observation

$\exists$ *optimal tree decomposition with taxa in decomposition leaves*

Dynamic Programming Idea

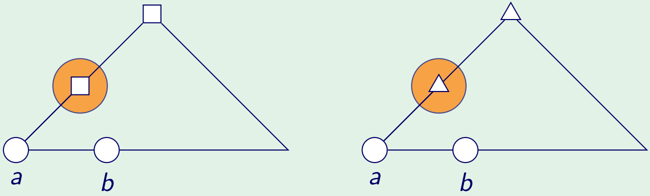


table: $[X, M \quad] =$ minimum #deletions "below" X respecting M

Ongoing Work: Towards a Practical Algorithm

Observation

$\exists$ *optimal tree decomposition with taxa in decomposition leaves*

Dynamic Programming Idea

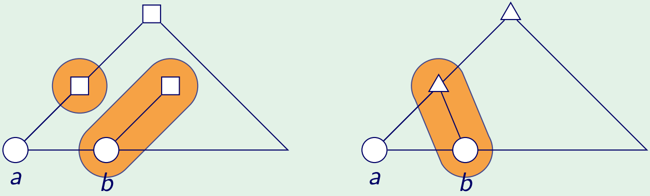


table: $[X, M \quad] =$ minimum #deletions "below" X respecting M

Ongoing Work: Towards a Practical Algorithm

Observation

∃ *optimal tree decomposition with taxa in decomposition leaves*

Dynamic Programming Idea

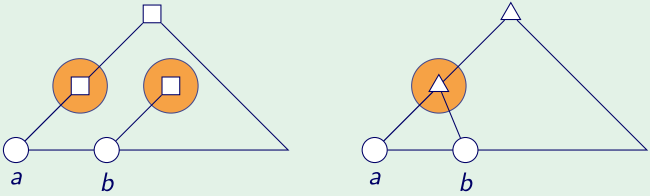


table: $[X, M \quad] =$ minimum #deletions “below” X respecting M

Ongoing Work: Towards a Practical Algorithm

Observation

$\exists$ *optimal tree decomposition with taxa in decomposition leaves*

Dynamic Programming Idea

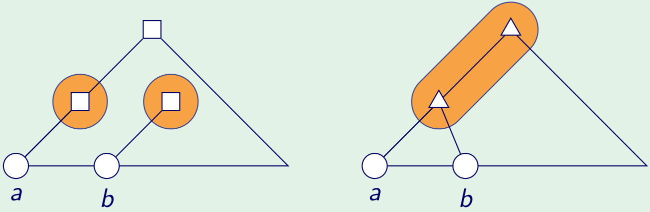


table: $[X, M \quad] =$ minimum #deletions "below" X respecting M

Ongoing Work: Towards a Practical Algorithm

Observation

∃ *optimal tree decomposition with taxa in decomposition leaves*

Dynamic Programming Idea

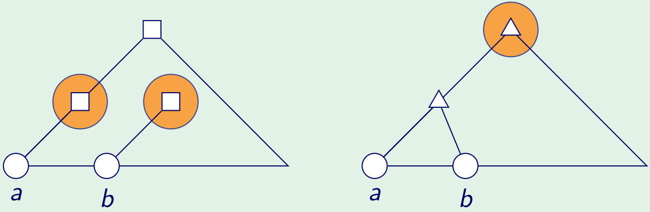


table: $[X, M \quad]$ = minimum #deletions "below" X respecting M

Ongoing Work: Towards a Practical Algorithm

Observation

$\exists$ *optimal tree decomposition with taxa in decomposition leaves*

Dynamic Programming Idea

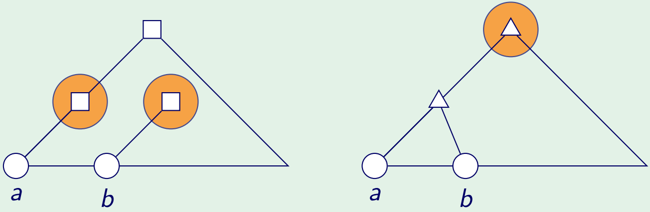


table: $[X, M, T]$ = minimum #deletions "below" X respecting M & T

Ongoing Work: Towards a Practical Algorithm

Observation

$\exists$ *optimal tree decomposition with taxa in decomposition leaves*

Dynamic Programming Idea

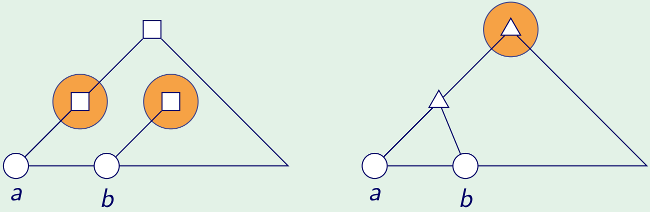


table: $[X, M, T]$ = minimum #deletions "below" X respecting M & T
$\rightsquigarrow O^*(\text{tw}^{\text{tw}})$ space

Ongoing Work: Towards a Practical Algorithm

Observation

$\exists$ *optimal tree decomposition with taxa in decomposition leaves*

Conjecture

$\exists$ optimal tree decomposition isomorphic to T_1 or T_2

Dynamic Programming Idea

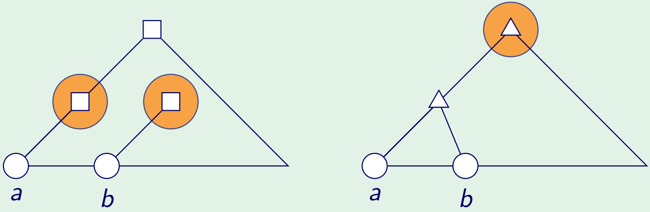


table: $[X, M, T]$ = minimum #deletions "below" X respecting M & T

$\rightsquigarrow O^*(\text{tw}^{\text{tw}})$ space

Thanks!