

PHYLOGENOMICS

EFFICIENT ANALYSIS BY ACCOUNTING FOR PHYLOGENETIC TERRACES

Chernomor O., Minh B.Q., and von Haeseler A.

University of Vienna, MFPL, CIBIV

13.06.2016

INPUT DATA

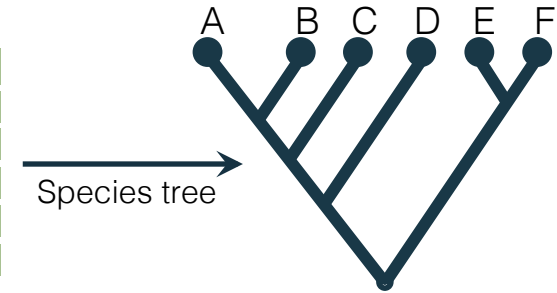
	Partition 1	Partition 2	Partition 3	Partition 4
A:	C-GTAAGCTGTAC	ATCCACA-GAGTA	GAG-ATACTGAT-	TCGA--GACATCA
B:	GAG-ATACTGTCA	ATTGAC-TGAGTA	CAGTAAG-TGTAC	TCAATTG-CATGA
C:	CTGTA--CTGCTC	CGT-CATGA-TAG	CA-TAAGCTGT-C	ACACGTGAC--GA
D:	CA-AAGCTG-CA-	CATGAC-TGAGTA	CTGTA--CTGTCA	CG-ATTGACATGA
E:	GTGTCA--TGTAC	ATG-ACAGGA-TG	CAGCA-GCTG-CA	TGCGTTG--TGGA
F:	CAGCA-GCTG-TA	TATG-CATGATCA	CA--AAGCTGTGC	TCAATCGACA-TA

➔ SUPERMATRIX

- Multiple sequence alignment consists of several blocks (genes / loci)

INPUT DATA

	Partition 1	Partition 2	Partition 3	Partition 4
A:	C-GTAAGCTGTAC	ATCCACA-GAGTA	GAG-ATACTGAT-	TCGA--GACATCA
B:	GAG-ATACTGTCA	ATTGAC-TGAGTA	CAGTAAG-TGTAC	TCAATTG-CATGA
C:	CTGTA--CTGCTC	CGT-CATGA-TAG	CA-TAAGCTGT-C	ACACGTGAC--GA
D:	CA-AAGCTG-CA-	CATGAC-TGAGTA	CTGTA--CTGTCA	CG-ATTGACATGA
E:	GTGTCA--TGTAC	ATG-ACAGGA-TG	CAGCA-GCTG-CA	TGCGTTG--TGGA
F:	CAGCA-GCTG-TA	TATG-CATGATCA	CA--AAGCTGTGC	TCAATCGACA-TA

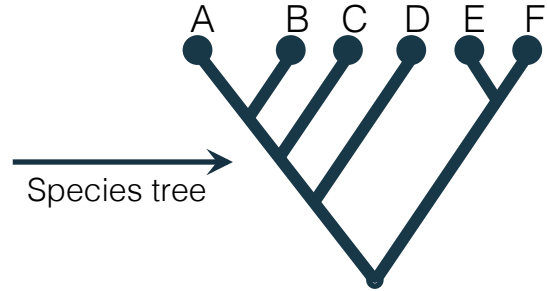


→ SUPERMATRIX

- Multiple sequence alignment consists of several blocks (genes / loci)

INPUT DATA

	Partition 1	Partition 2	Partition 3	Partition 4
A:	C-GTAAGCTGTAC	ATCCACA-GAGTA	GAG-ATACTGAT-	-----
B:	GAG-ATACTGTCA	-----	CAGTAAG-TGTAC	-----
C:	CTGTA--CTGCTC	-----	-----	ACACGTGAC--GA
D:	-----	CATGAC-TGAGTA	CTGTA--CTGTCA	CG-ATTGACATGA
E:	-----	ATG-ACAGGA-TG	-----	TGCGTTG--TGGA
F:	CAGCA-GCTG-TA	TATG-CATGATCA	CA--AAGCTGTGC	TCAATCGACA-TA



→ SUPERMATRIX

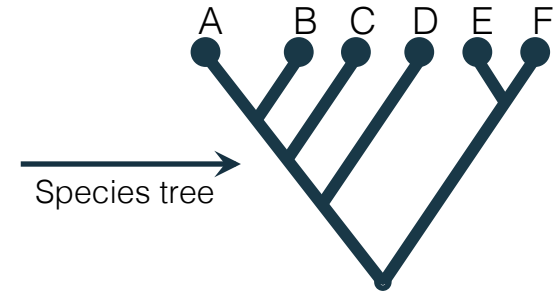
- Multiple sequence alignment consists of several blocks (genes / loci)

→ MISSING SEQUENCES (not sequenced / low quality / gene is absent in some species)

- Sequence of gaps

INPUT DATA

	Partition 1	Partition 2	Partition 3	Partition 4
A:	C-GTAAGCTGTAC	ATCCACA-GAGTA	GAG-ATACTGAT-	-----
B:	GAG-ATACTGTCA	-----	CAGTAAG-TGTAC	-----
C:	CTGTA--CTGCTC	-----	-----	ACACGTGAC--GA
D:	-----	CATGAC-TGAGTA	CTGTA--CTGTCA	CG-ATTGACATGA
E:	-----	ATG-ACAGGA-TG	-----	TGCGTTG--TGGA
F:	CAGCA-GCTG-TA	TATG-CATGATCA	CA--AAGCTGTGC	TCAATCGACA-TA



→ SUPERMATRIX

- Multiple sequence alignment consists of several blocks (genes / loci)

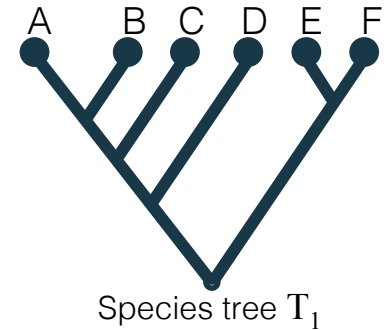
→ MISSING SEQUENCES (not sequenced / low quality / gene is absent in some species)

- Sequence of gaps
- Missing data

SPECIES TREE AND INDUCED PARTITION TREES

Partition 1

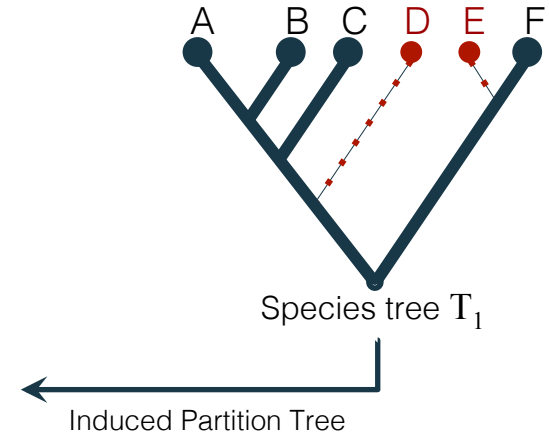
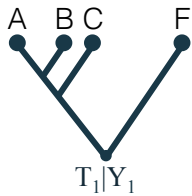
A:	C-GTAAGCTGTAC
B:	GAG-ATACTGTCA
C:	CTGTA--CTGCTC
D:	-----
E:	-----
F:	CAGCA-GCTG-TA



SPECIES TREE AND INDUCED PARTITION TREES

Partition 1

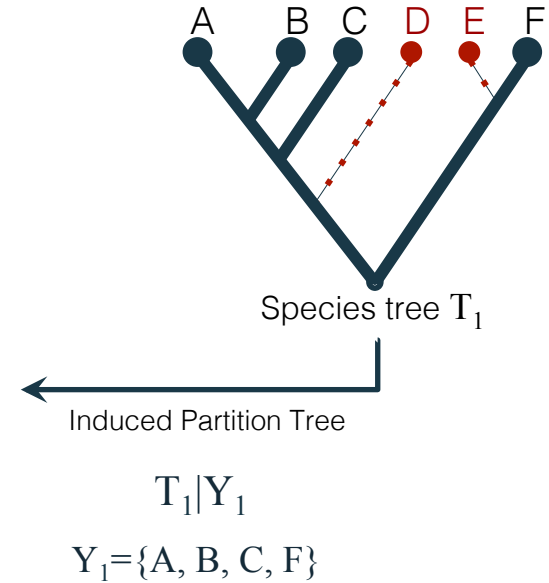
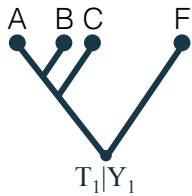
A:	C-GTAAGCTGTAC
B:	GAG-ATACTGTCA
C:	CTGTA--CTGCTC
D:	-----
E:	-----
F:	CAGCA-GCTG-TA



SPECIES TREE AND INDUCED PARTITION TREES

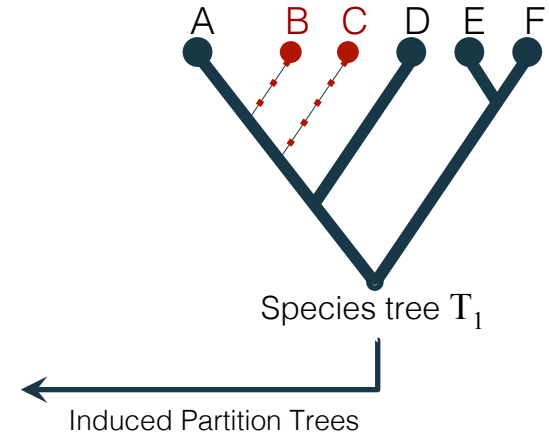
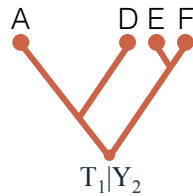
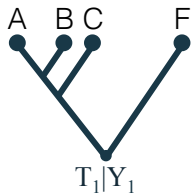
Partition 1

A:	C-GTAAGCTGTAC
B:	GAG-ATACTGTCA
C:	CTGTA--CTGCTC
D:	-----
E:	-----
F:	CAGCA-GCTG-TA



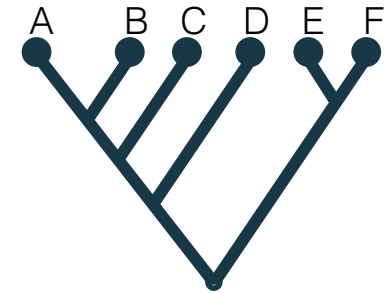
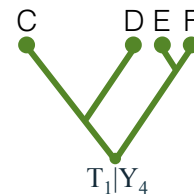
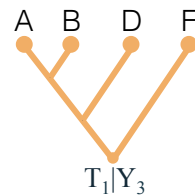
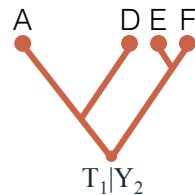
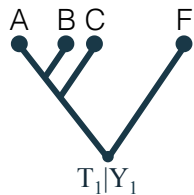
SPECIES TREE AND INDUCED PARTITION TREES

	Partition 1	Partition 2
A:	C-GTAAGCTGTAC	ATCCACA-GAGTA
B:	GAG-ATACTGTCA	-----
C:	CTGTA--CTGCTC	-----
D:	-----	CATGAC-TGAGTA
E:	-----	ATG-ACAGGA-TG
F:	CAGCA-GCTG-TA	TATG-CATGATCA



SPECIES TREE AND INDUCED PARTITION TREES

	Partition 1	Partition 2	Partition 3	Partition 4
A:	C-GTAAGCTGTAC	ATCCACA-GAGTA	GAG-ATACTGAT-	-----
B:	GAG-ATACTGTCA	-----	CAGTAAG-TGTAC	-----
C:	CTGTA--CTGCTC	-----	-----	ACACGTGAC--GA
D:	-----	CATGAC-TGAGTA	CTGTA--CTGTCA	CG-ATTGACATGA
E:	-----	ATG-ACAGGA-TG	-----	TGCGTTG--TGGA
F:	CAGCA-GCTG-TA	TATG-CATGATCA	CA--AAGCTGTGC	TCAATCGACA-TA

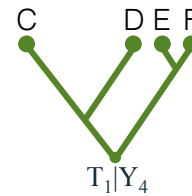
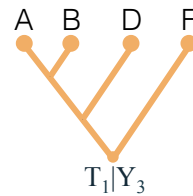
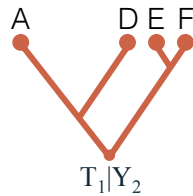
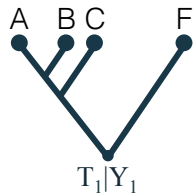


Species tree T_1

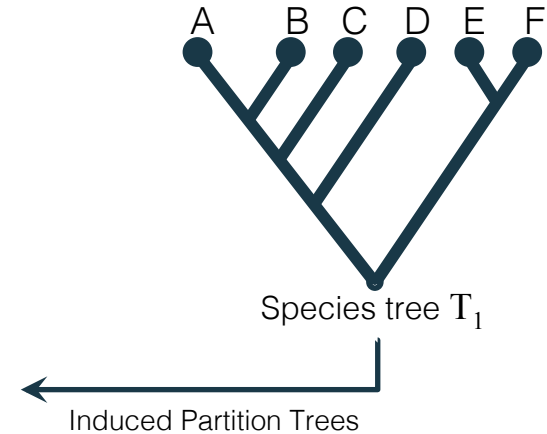
Induced Partition Trees

SPECIES TREE AND INDUCED PARTITION TREES

	Partition 1	Partition 2	Partition 3	Partition 4
A:	C-GTAAGCTGTAC	ATCCACA-GAGTA	GAG-ATACTGAT-	-----
B:	GAG-ATACTGTCA	-----	CAGTAAG-TGTAC	-----
C:	CTGTA--CTGCTC	-----	-----	ACACGTGAC--GA
D:	-----	CATGAC-TGAGTA	CTGTA--CTGTCA	CG-ATTGACATGA
E:	-----	ATG-ACAGGA-TG	-----	TGCGTTG--TGGA
F:	CAGCA-GCTG-TA	TATG-CATGATCA	CA--AAGCTGTGC	TCAATCGACA-TA

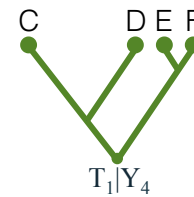
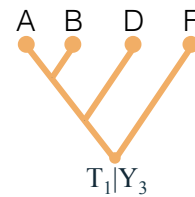
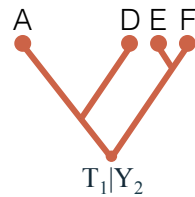
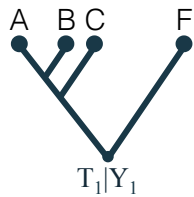


$$L(T_1) = L_1 + L_2 + L_3 + L_4$$

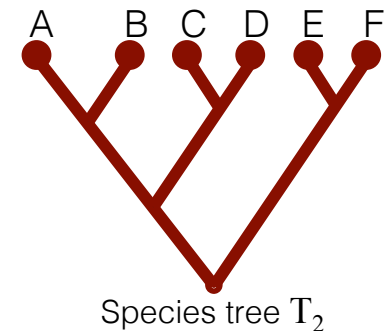
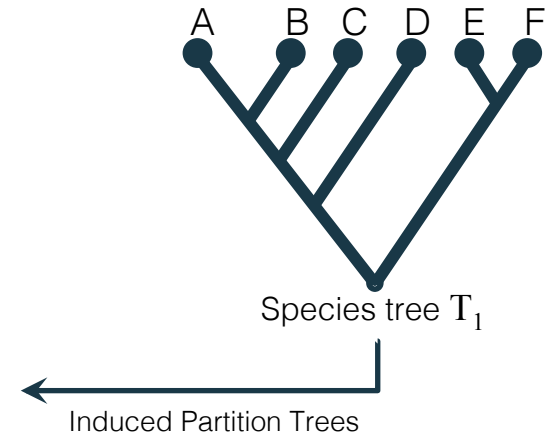


THE CONCEPT OF PHYLOGENETIC TERRACES

	Partition 1	Partition 2	Partition 3	Partition 4
A:	C-GTAAGCTGTAC	ATCCACA-GAGTA	GAG-ATACTGAT-	-----
B:	GAG-ATACTGTCA	-----	CAGTAAG-TGTAC	-----
C:	CTGTA--CTGCTC	-----	-----	ACACGTGAC--GA
D:	-----	CATGAC-TGAGTA	CTGTA--CTGTCA	CG-ATTGACATGA
E:	-----	ATG-ACAGGA-TG	-----	TGCGTTG--TGGA
F:	CAGCA-GCTG-TA	TATG-CATGATCA	CA--AAGCTGTGC	TCAATCGACA-TA

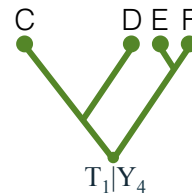
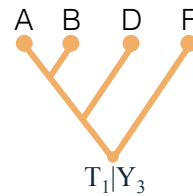
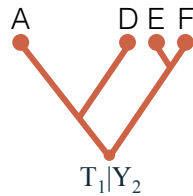
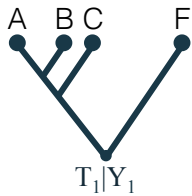


$$L(T_1) = L_1 + L_2 + L_3 + L_4$$

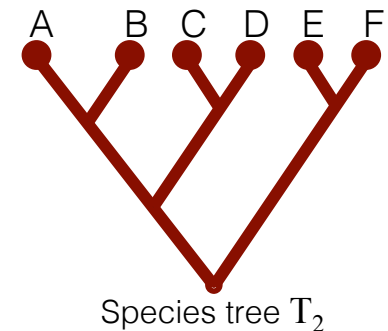
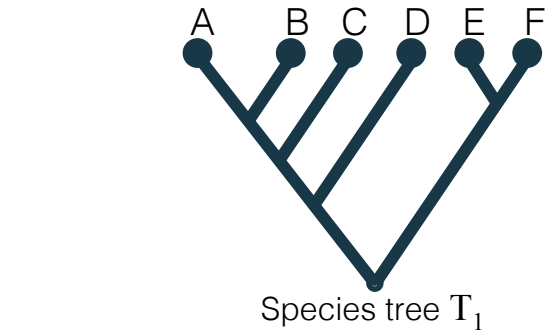


THE CONCEPT OF PHYLOGENETIC TERRACES

	Partition 1	Partition 2	Partition 3	Partition 4
A:	C-GTAAGCTGTAC	ATCCACA-GAGTA	GAG-ATACTGAT-	-----
B:	GAG-ATACTGTCA	-----	CAGTAAG-TGTAC	-----
C:	CTGTA--CTGCTC	-----	-----	ACACGTGAC--GA
D:	-----	CATGAC-TGAGTA	CTGTA--CTGTCA	CG-ATTGACATGA
E:	-----	ATG-ACAGGA-TG	-----	TGCGTTG--TGGA
F:	CAGCA-GCTG-TA	TATG-CATGATCA	CA--AAGCTGTGC	TCAATCGACA-TA

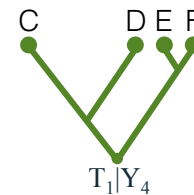
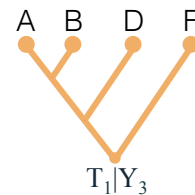
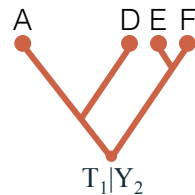
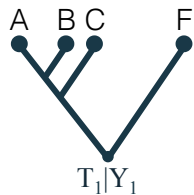


$$L(T_1) = L_1 + L_2 + L_3 + L_4$$

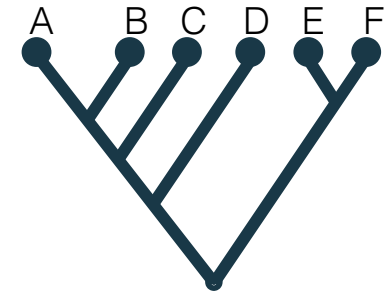


THE CONCEPT OF PHYLOGENETIC TERRACES

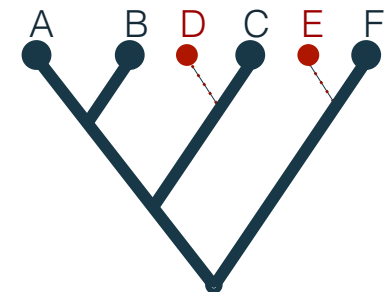
	Partition 1	Partition 2	Partition 3	Partition 4
A:	C-GTAAGCTGTAC	ATCCACA-GAGTA	GAG-ATACTGAT-	-----
B:	GAG-ATACTGTCA	-----	CAGTAAG-TGTAC	-----
C:	CTGTA--CTGCTC	-----	-----	ACACGTGAC--GA
D:	-----	CATGAC-TGAGTA	CTGTA--CTGTCA	CG-ATTGACATGA
E:	-----	ATG-ACAGGA-TG	-----	TGCGTTG--TGGA
F:	CAGCA-GCTG-TA	TATG-CATGATCA	CA--AAGCTGTGC	TCAATCGACA-TA



$$L(T_1) = L_1 + L_2 + L_3 + L_4$$



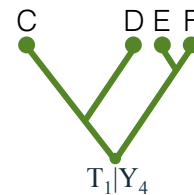
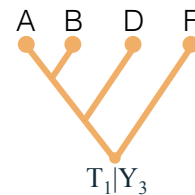
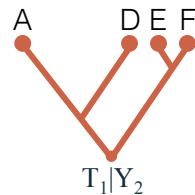
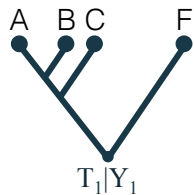
Species tree T_1



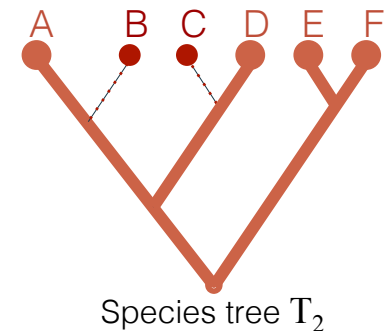
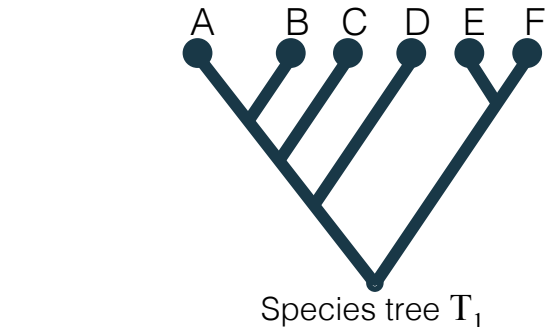
Species tree T_2

THE CONCEPT OF PHYLOGENETIC TERRACES

	Partition 1	Partition 2	Partition 3	Partition 4
A:	C-GTAAGCTGTAC	ATCCACA-GAGTA	GAG-ATACTGAT-	-----
B:	GAG-ATACTGTCA	-----	CAGTAAG-TGTAC	-----
C:	CTGTA--CTGCTC	-----	-----	ACACGTGAC--GA
D:	-----	CATGAC-TGAGTA	CTGTA--CTGTCA	CG-ATTGACATGA
E:	-----	ATG-ACAGGA-TG	-----	TGCGTTG--TGGA
F:	CAGCA-GCTG-TA	TATG-CATGATCA	CA--AAGCTGTGC	TCAATCGACA-TA

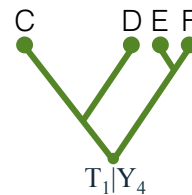
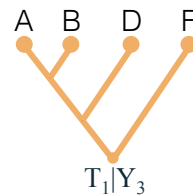
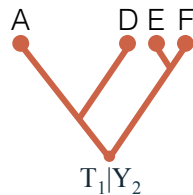
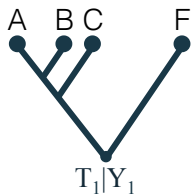


$$L(T_1) = L_1 + L_2 + L_3 + L_4$$



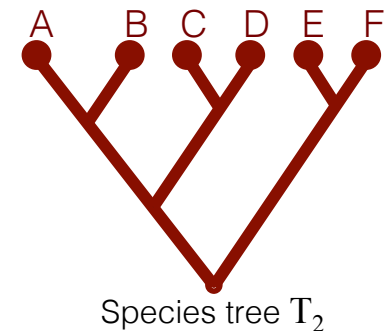
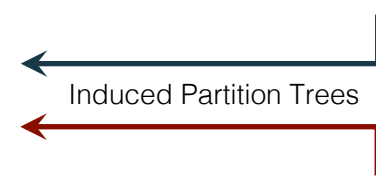
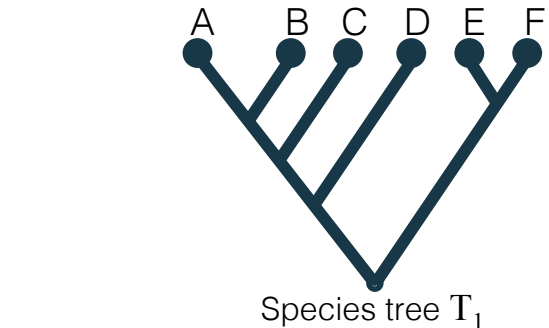
THE CONCEPT OF PHYLOGENETIC TERRACES

	Partition 1	Partition 2	Partition 3	Partition 4
A:	C-GTAAGCTGTAC	ATCCACA-GAGTA	GAG-ATACTGAT-	-----
B:	GAG-ATACTGTCA	-----	CAGTAAG-TGTAC	-----
C:	CTGTA--CTGCTC	-----	-----	ACACGTGAC--GA
D:	-----	CATGAC-TGAGTA	CTGTA--CTGTCA	CG-ATTGACATGA
E:	-----	ATG-ACAGGA-TG	-----	TGCGTTG--TGGA
F:	CAGCA-GCTG-TA	TATG-CATGATCA	CA--AAGCTGTGC	TCAATCGACA-TA



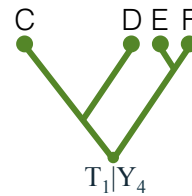
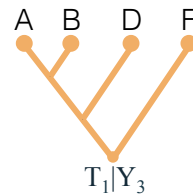
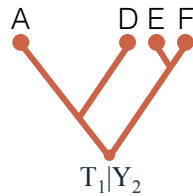
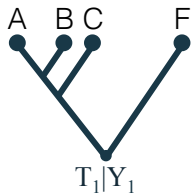
$$L(T_1) = L_1 + L_2 + L_3 + L_4$$

$$L(T_2) = L_1 + L_2 + L_3 + L_4$$

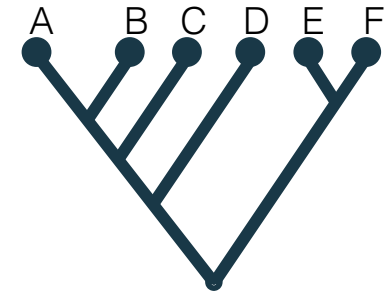


THE CONCEPT OF PHYLOGENETIC TERRACES

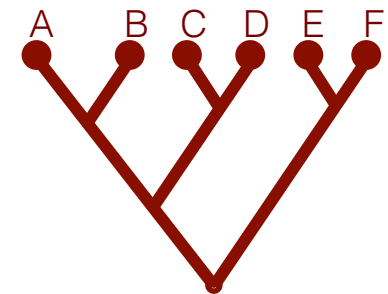
	Partition 1	Partition 2	Partition 3	Partition 4
A:	C-GTAAGCTGTAC	ATCCACA-GAGTA	GAG-ATACTGAT-	-----
B:	GAG-ATACTGTCA	-----	CAGTAAG-TGTAC	-----
C:	CTGTA--CTGCTC	-----	-----	ACACGTGAC--GA
D:	-----	CATGAC-TGAGTA	CTGTA--CTGTCA	CG-ATTGACATGA
E:	-----	ATG-ACAGGA-TG	-----	TGCGTTG--TGGA
F:	CAGCA-GCTG-TA	TATG-CATGATCA	CA--AAGCTGTGC	TCAATCGACA-TA



$$L(T_1) = L(T_2)$$



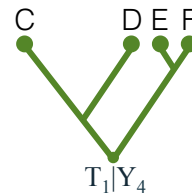
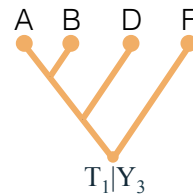
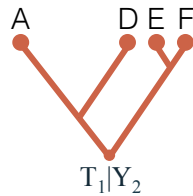
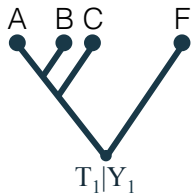
Species tree T_1



Species tree T_2

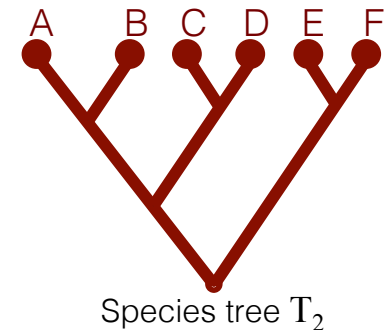
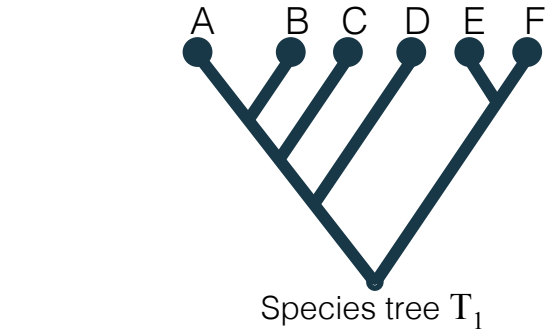
THE CONCEPT OF PHYLOGENETIC TERRACES

	Partition 1	Partition 2	Partition 3	Partition 4
A:	C-GTAAGCTGTAC	ATCCACA-GAGTA	GAG-ATACTGAT-	-----
B:	GAG-ATACTGTCA	-----	CAGTAAG-TGTAC	-----
C:	CTGTA--CTGCTC	-----	-----	ACACGTGAC--GA
D:	-----	CATGAC-TGAGTA	CTGTA--CTGTCA	CG-ATTGACATGA
E:	-----	ATG-ACAGGA-TG	-----	TGCGTTG--TGGA
F:	CAGCA-GCTG-TA	TATG-CATGATCA	CA--AAGCTGTGC	TCAATCGACA-TA



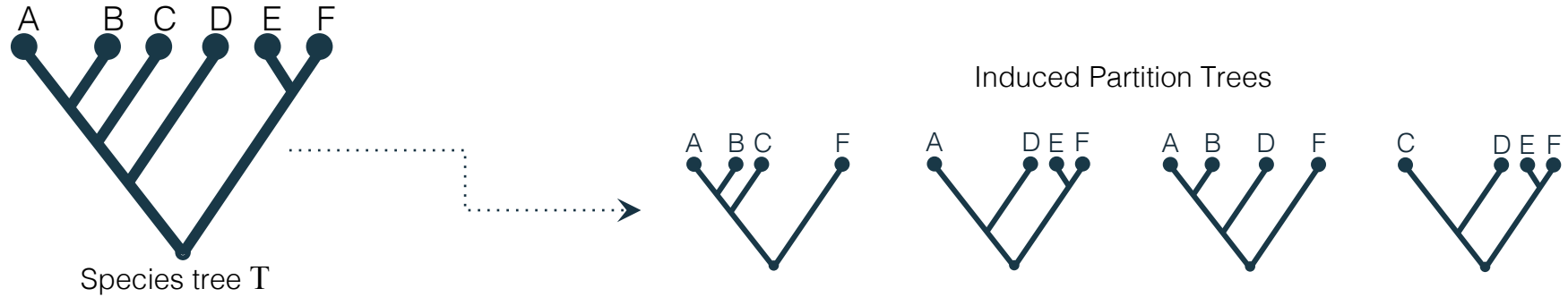
$$L(T_1) = L(T_2)$$

It is said that T_1 and T_2 belong to one
Phylogenetic Terrace⁽¹⁾.

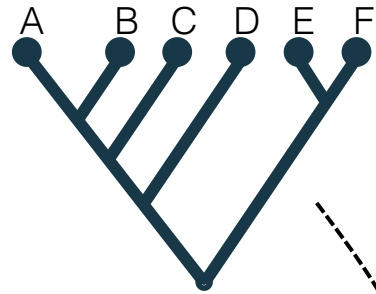


⁽¹⁾ **M.J. Sanderson et al.** Terraces in Phylogenetic Tree Space. *Science*. **2011**: 333 (6041), 448-450.

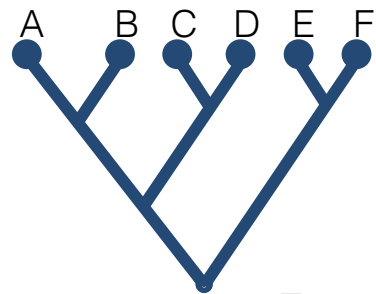
DETECTING TERRACES DURING TREE SEARCH



DETECTING TERRACES DURING TREE SEARCH

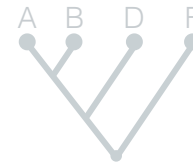


Species tree T

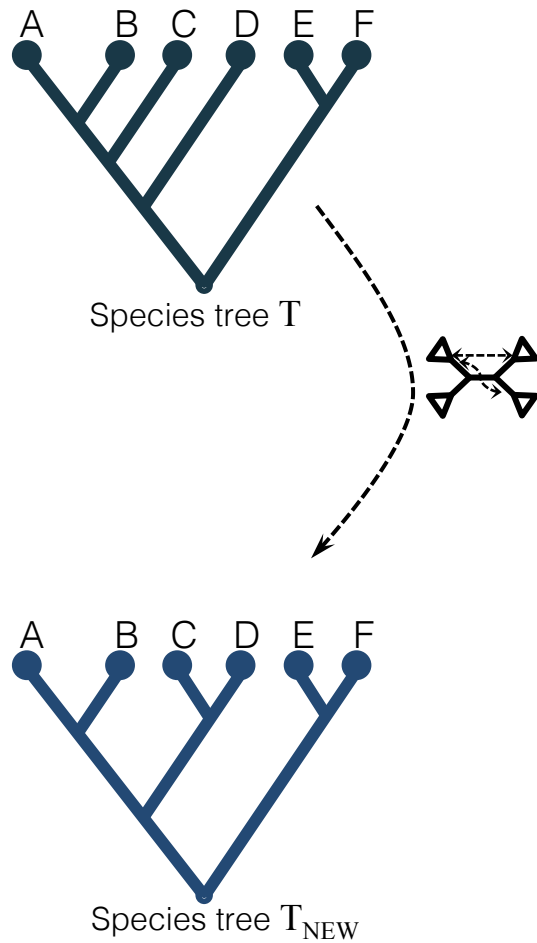


Species tree T_{NEW}

Induced Partition Trees



DETECTING TERRACES DURING TREE SEARCH



Induced Partition Trees



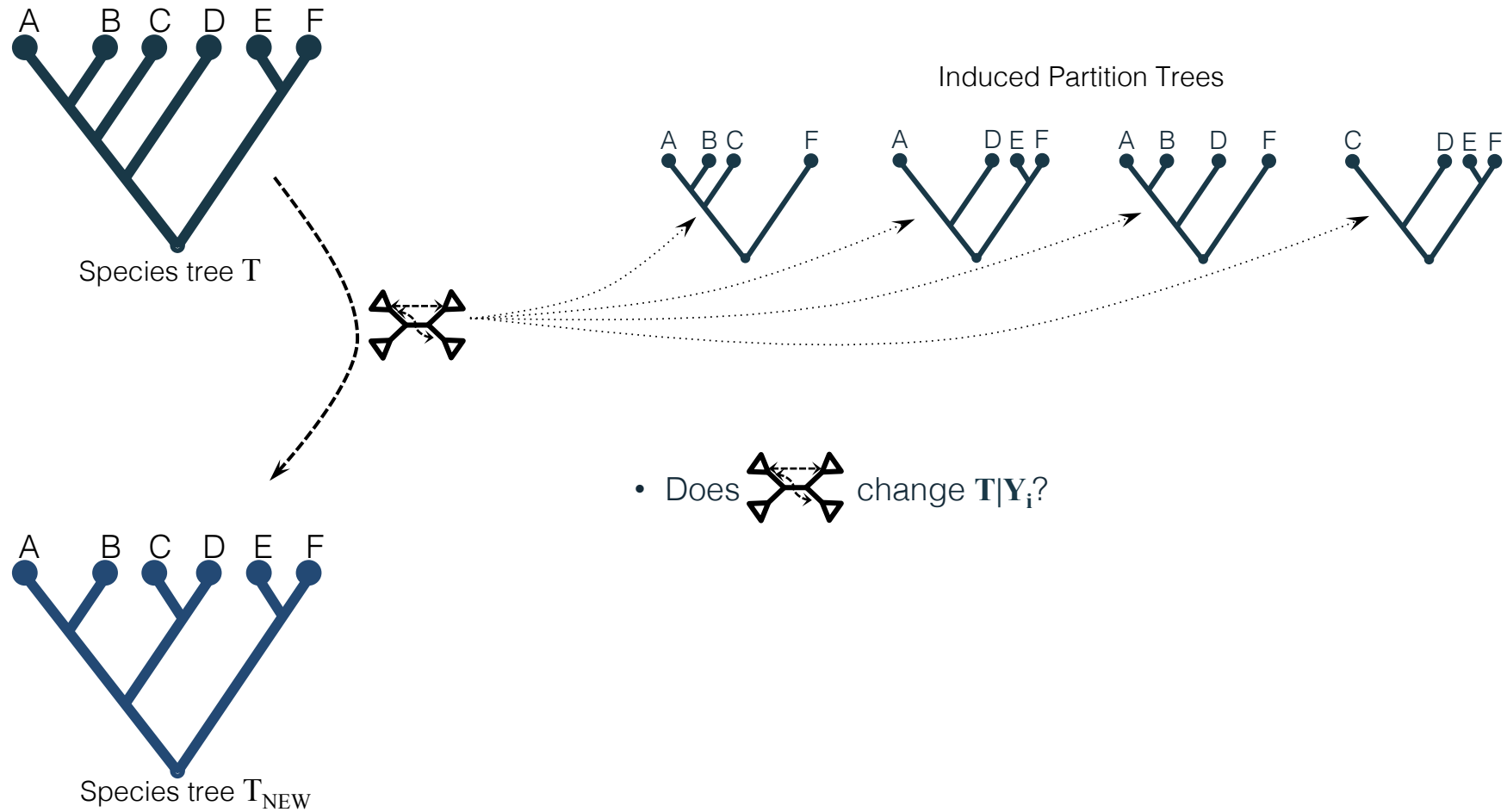
- Topological Rearrangements**

NNI

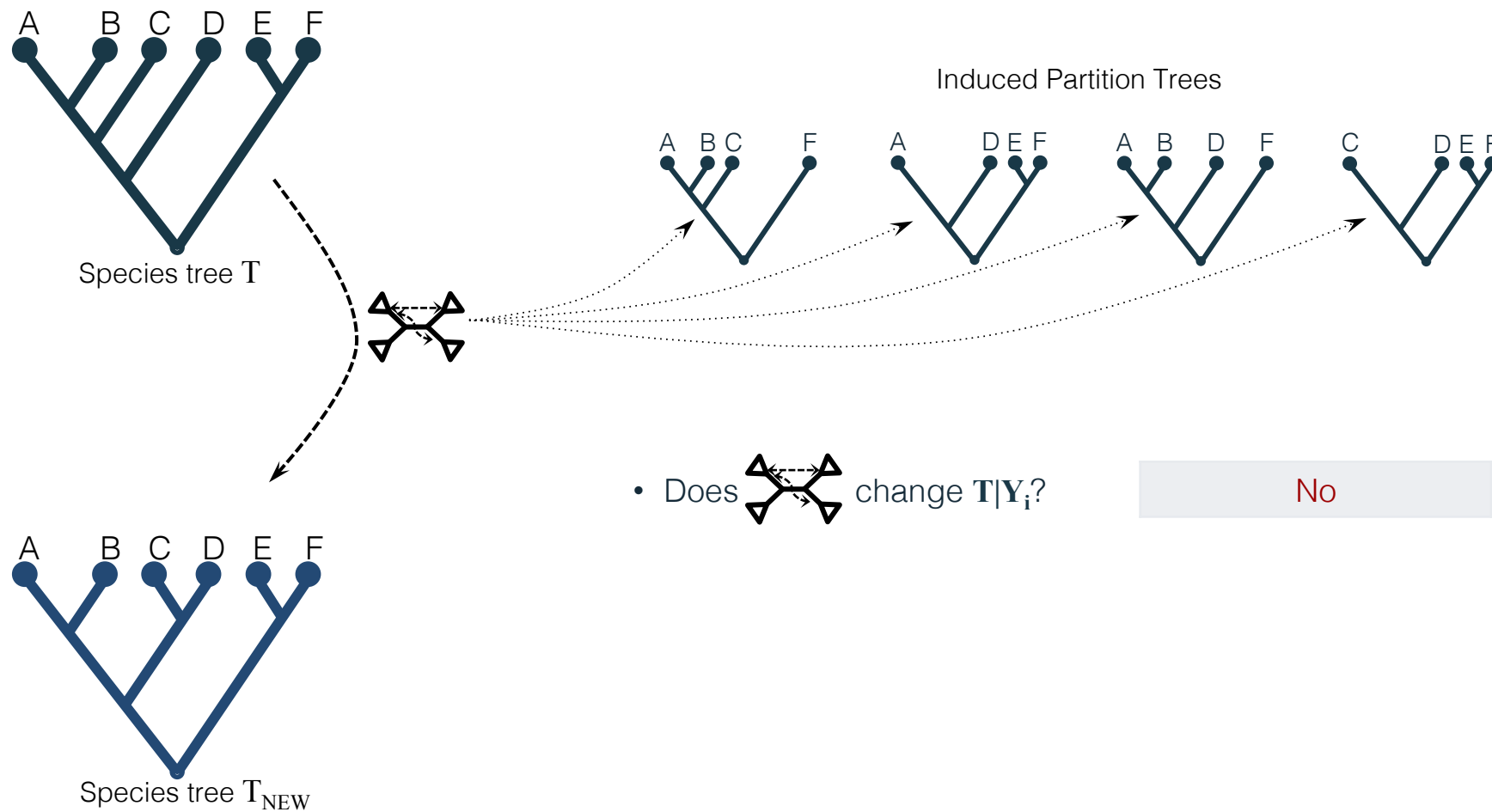
SPR

TBR

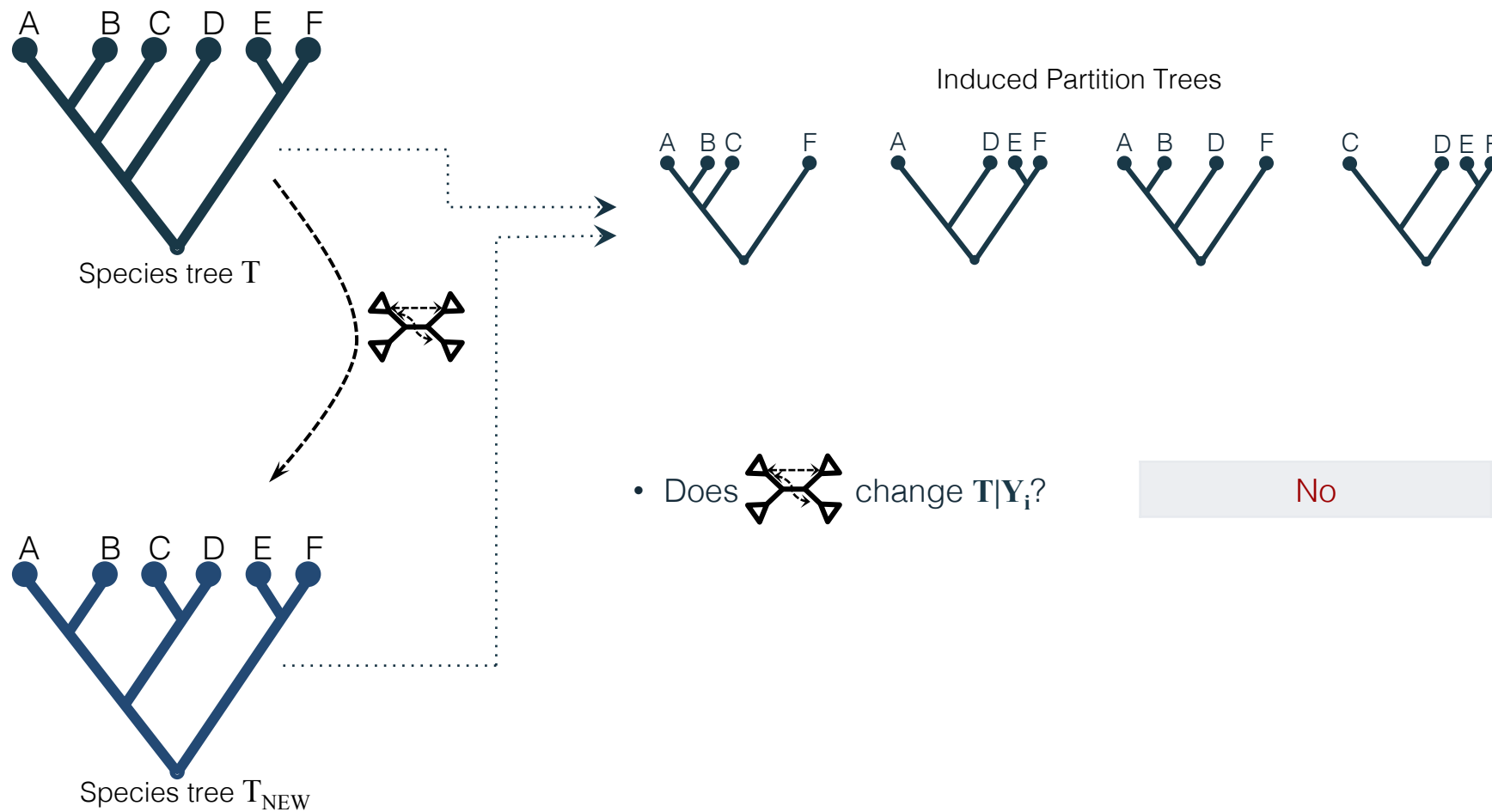
DETECTING TERRACES DURING TREE SEARCH



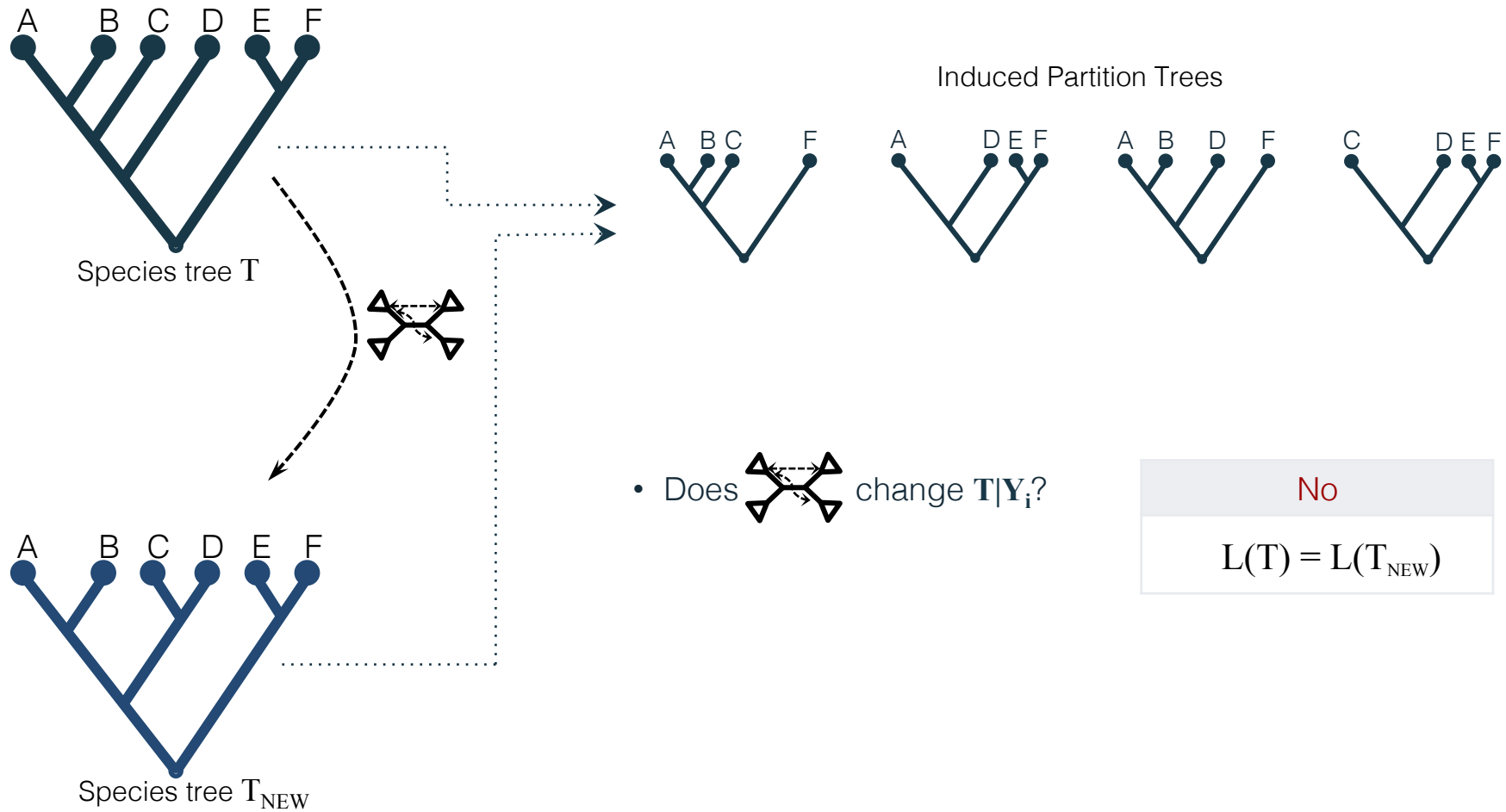
DETECTING TERRACES DURING TREE SEARCH



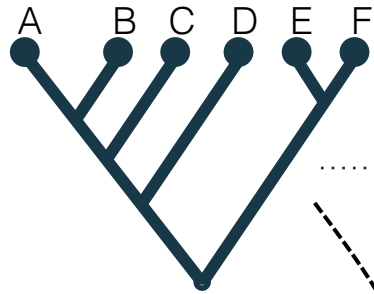
DETECTING TERRACES DURING TREE SEARCH



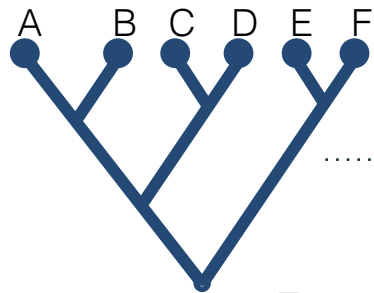
DETECTING TERRACES DURING TREE SEARCH



DETECTING TERRACES DURING TREE SEARCH

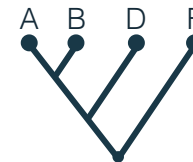
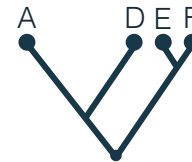
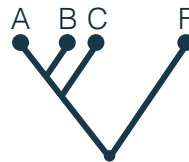


Species tree T



Species tree T_{NEW}

Induced Partition Trees



- Does  change $T|Y_i$?

No

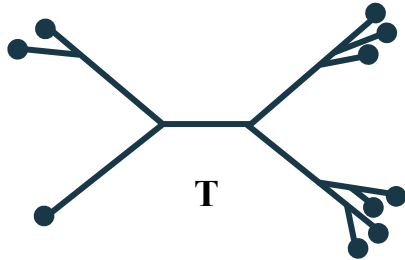
$$L(T) = \cancel{L(T_{\text{NEW}})}$$

save time



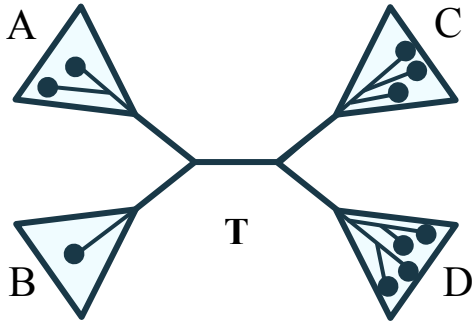
TOPOLOGICAL REARRANGEMENT ON SPECIES TREE

- NEAREST NEIGHBOR INTERCHANGED (NNI)



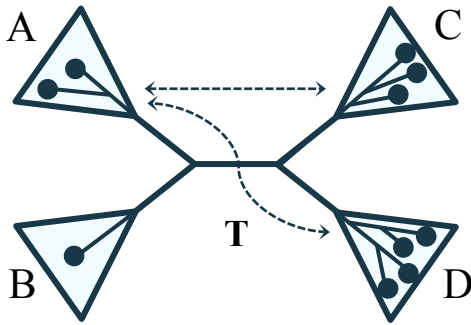
TOPOLOGICAL REARRANGEMENT ON SPECIES TREE

- NEAREST NEIGHBOR INTERCHANGED (NNI)



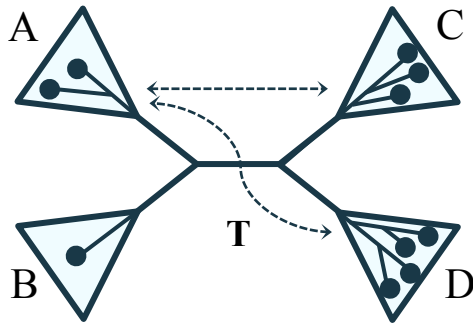
TOPOLOGICAL REARRANGEMENT ON SPECIES TREE

- NEAREST NEIGHBOR INTERCHANGED (NNI)

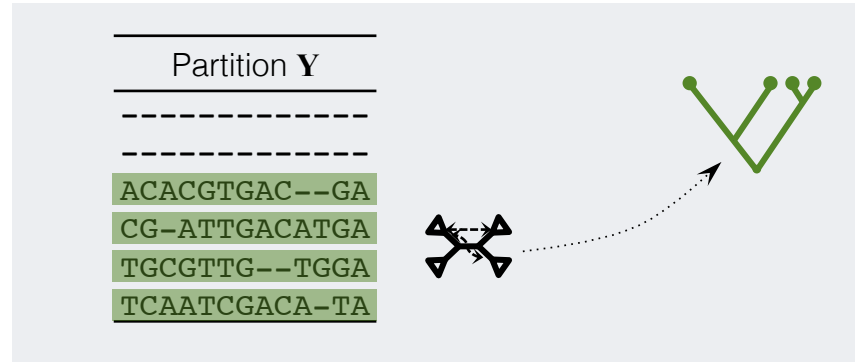


NNI ON SPECIES TREE AND ITS EFFECT ON PARTITION TREE

- NEAREST NEIGHBOR INTERCHANGED (NNI)

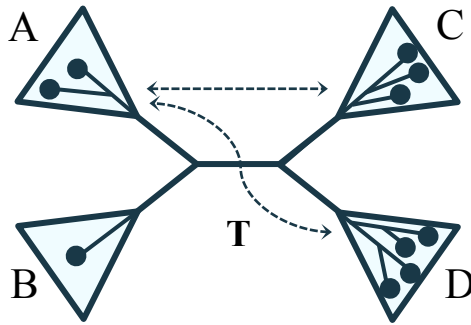


Species Tree

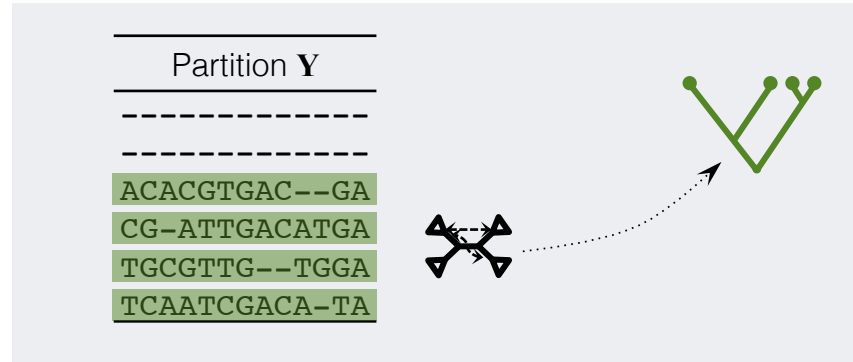


NNI ON SPECIES TREE AND ITS EFFECT ON PARTITION TREE

- NEAREST NEIGHOR INTERCHANGED (NNI)



Species Tree

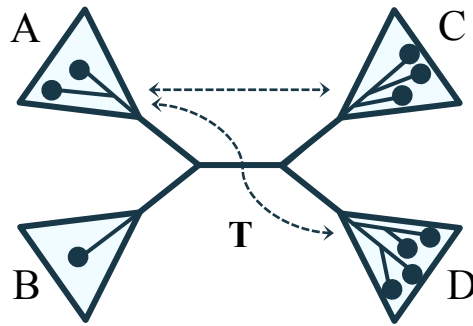


INTUITION

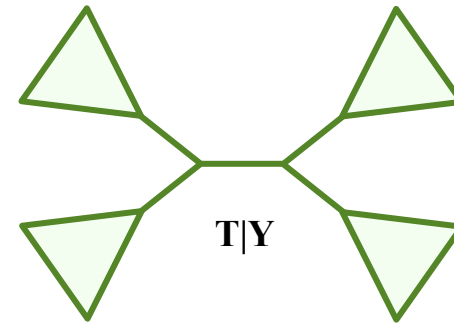
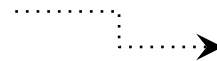
NNI on Species Tree
should correspond to
an NNI on Induced
Partition Tree.

NNI ON SPECIES TREE AND ITS EFFECT ON PARTITION TREE

- NEAREST NEIGHOR INTERCHANGED (NNI)



Species Tree



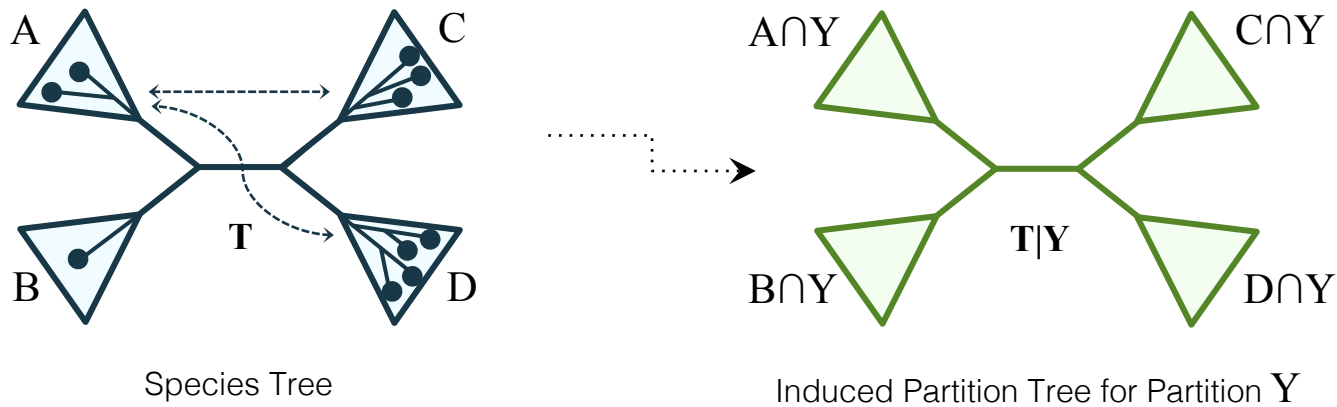
Induced Partition Tree for Partition Y

INTUITION

NNI on Species Tree
should correspond to
an NNI on Induced
Partition Tree.

NNI ON SPECIES TREE AND ITS EFFECT ON PARTITION TREE

- NEAREST NEIGHOR INTERCHANGED (NNI)

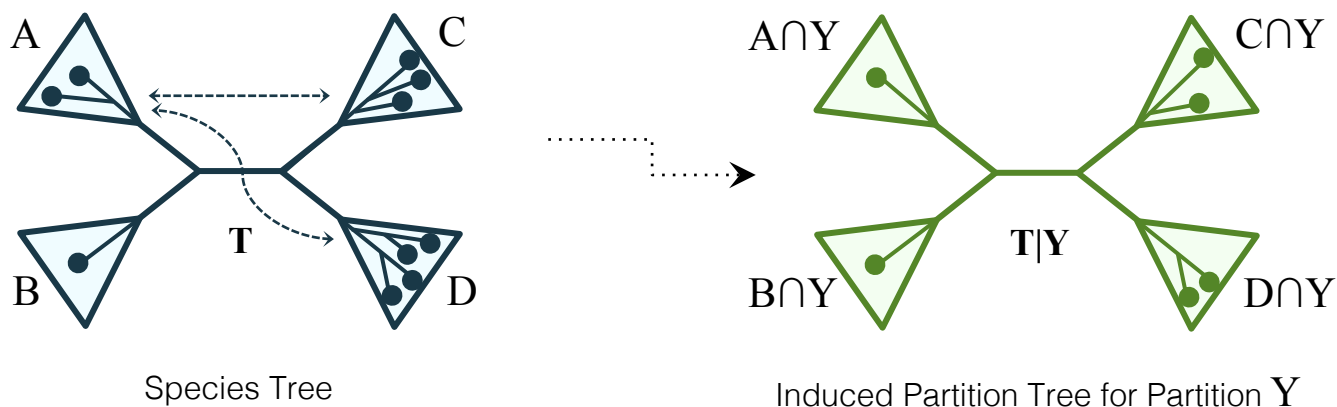


INTUITION

NNI on Species Tree
should correspond to
an NNI on Induced
Partition Tree.

NNI ON SPECIES TREE AND ITS EFFECT ON PARTITION TREE

- NEAREST NEIGHOR INTERCHANGED (NNI)



INTUITION

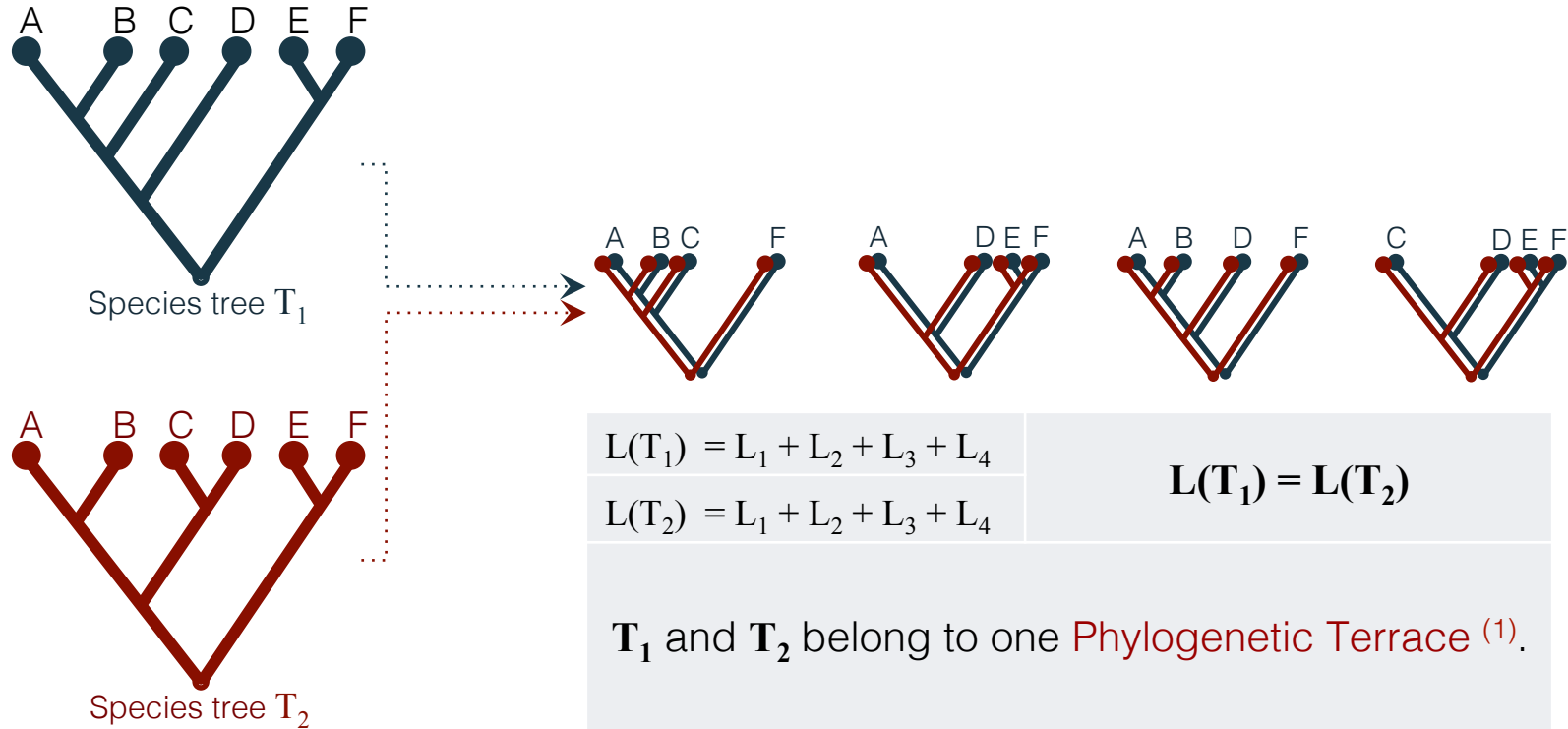
NNI on Species Tree
should correspond to
an NNI on Induced
Partition Tree.

PROPOSITION 1: ⁽²⁾

For a given partition Y
all $A \cap Y$, $B \cap Y$, $C \cap Y$,
 $D \cap Y$ must be non-
empty.

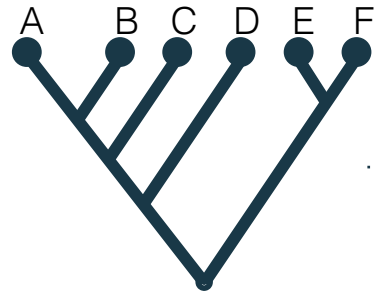
⁽²⁾ O. Chernomor et al. 2015. Consequences of Common Topological Rearrangements for Partition Trees in Phylogenomic Inference. J Comput Biol 22(12):1129-42.

RECAP: PHYLOGENETIC TERRACES

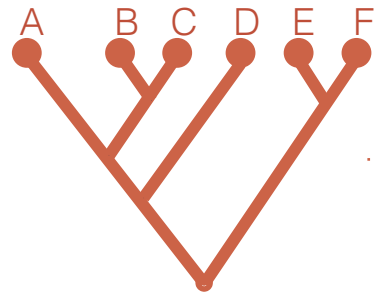
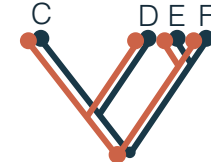
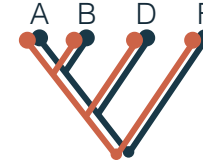
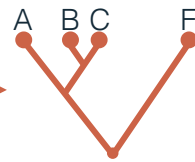
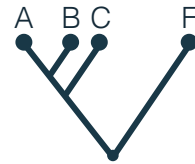


⁽¹⁾ **M.J. Sanderson et al.** Terraces in Phylogenetic Tree Space. *Science*. **2011**: 333 (6041), 448-450.

THE CONCEPT OF PARTIAL TERRACES



Species tree T_1



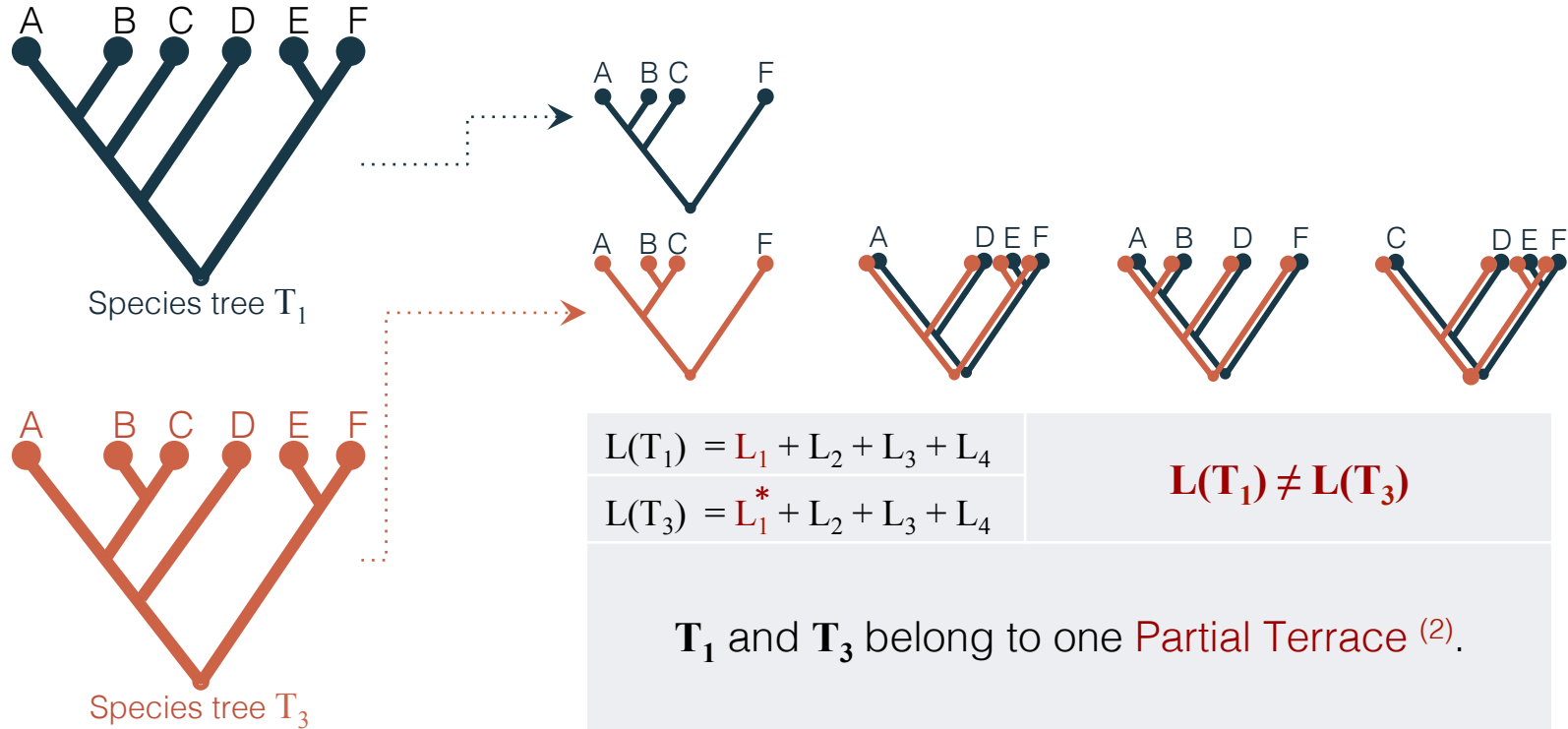
Species tree T_3

$$L(T_1) = L_1 + L_2 + L_3 + L_4$$

$$L(T_3) = L_1^* + L_2 + L_3 + L_4$$

$$L(T_1) \neq L(T_3)$$

THE CONCEPT OF PARTIAL TERRACES



⁽²⁾ O. Chernomor et al. 2015. Consequences of Common Topological Rearrangements for Partition Trees in Phylogenomic Inference.

TERRACE-AWARE MAXIMUM LIKELIHOOD INFERENCE

- APPLICATION

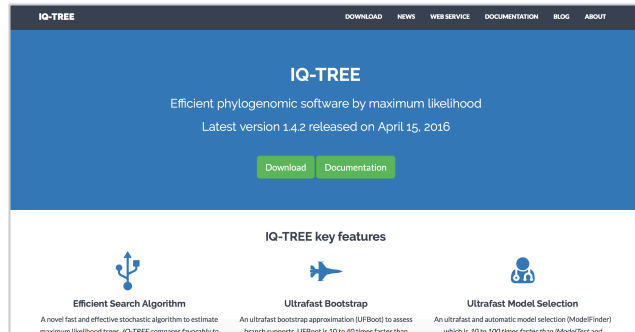
- ➔ Maximum Likelihood Inference

TERRACE-AWARE MAXIMUM LIKELIHOOD INFERENCE

- APPLICATION

➔ Maximum Likelihood Inference

IQ-TREE



TERRACE-AWARE MAXIMUM LIKELIHOOD INFERENCE

- APPLICATION

➔ Maximum Likelihood Inference

IQ-TREE

The screenshot displays the IQ-TREE website. At the top, a navigation bar includes links for DOWNLOAD, NEWS, WEB SERVICE, DOCUMENTATION, BLOG, and ABOUT. The main header area is blue and contains the text 'IQ-TREE', 'Efficient phylogenomic software by maximum likelihood', and 'Latest version 1.4.2 released on April 15, 2016'. Below this are 'Download' and 'Documentation' buttons. The 'IQ-TREE key features' section highlights three features: 'Efficient Search Algorithm' (described as a novel fast and effective stochastic algorithm), 'Ultrafast Bootstrap' (described as an ultrafast bootstrap approximation), and 'Ultrafast Model Selection' (described as an ultrafast and automatic model selection). The 'Versions' section is divided into two columns: the left column lists 'IQ-TREE' in green, and the right column lists 'IQ-TREE_{PTA}' in red, with 'Terraces' and 'Partial Terraces' listed below it.

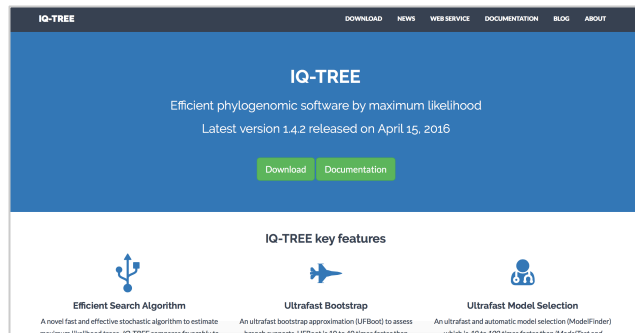
Versions	
IQ-TREE	IQ-TREE _{PTA}
	Terraces Partial Terraces

TERRACE-AWARE MAXIMUM LIKELIHOOD INFERENCE

• APPLICATION

➔ Maximum Likelihood Inference

IQ-TREE



The screenshot shows the IQ-TREE website. At the top, there is a navigation bar with links: DOWNLOAD, NEWS, WEB SERVICE, DOCUMENTATION, BLOG, and ABOUT. The main header area is blue and contains the text "IQ-TREE", "Efficient phylogenomic software by maximum likelihood", and "Latest version 1.4.2 released on April 15, 2016". Below this are two green buttons: "Download" and "Documentation". The "IQ-TREE key features" section is below, featuring three icons and their descriptions: "Efficient Search Algorithm" (a USB icon), "Ultrafast Bootstrap" (a jet icon), and "Ultrafast Model Selection" (a person icon). Each feature has a brief description of its capabilities.

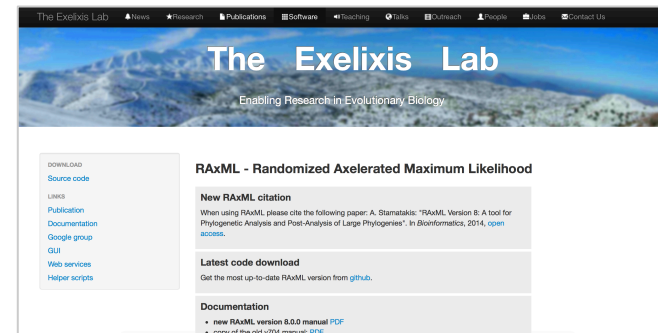
Versions

IQ-TREE

IQ-TREE_{PTA}

Terraces
Partial Terraces

RAxML

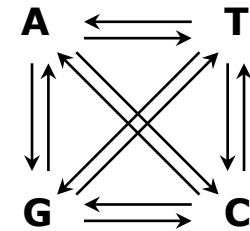


The screenshot shows the RAxML website. At the top, there is a navigation bar with links: The Exelixis Lab, News, Research, Publications, Software, Teaching, Talks, Outreach, People, Jobs, and Contact Us. The main header area features a landscape image and the text "The Exelixis Lab" and "Enabling Research in Evolutionary Biology". Below this is a sidebar with links: DOWNLOAD, Source code, LINKS, Publication, Documentation, Google group, Git, Web services, and Helper scripts. The main content area is titled "RAxML - Randomized Axelerated Maximum Likelihood" and contains sections for "New RAxML citation", "Latest code download", and "Documentation".

PARTITION MODELS IN MAXIMUM LIKELIHOOD (ML)

	Partition 1	Partition 2	Partition 3	Partition 4
A:	C-GTAAGCTGTAC	ATCCACA-GAGTA	GAG-ATACTGAT-	-----
B:	GAG-ATACTGTCA	-----	CAGTAAG-TGTAC	-----
C:	CTGTA--CTGCTC	-----	-----	ACACGTGAC--GA
D:	-----	CATGAC-TGAGTA	CTGTA--CTGTCA	CG-ATTGACATGA
E:	-----	ATG-ACAGGA-TG	-----	TGCGTTG--TGGA
F:	CAGCA-GCTG-TA	TATG-CATGATCA	CA--AAGCTGTGC	TCAATCGACA-TA

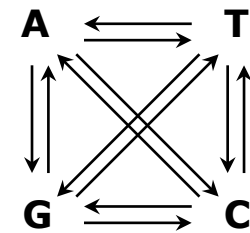
Substitution Model



PARTITION MODELS IN MAXIMUM LIKELIHOOD (ML)

	Partition 1	Partition 2	Partition 3	Partition 4
A:	C-GTAAGCTGTAC	ATCCACA-GAGTA	GAG-ATACTGAT-	-----
B:	GAG-ATACTGTCA	-----	CAGTAAG-TGTAC	-----
C:	CTGTA--CTGCTC	-----	-----	ACACGTGAC--GA
D:	-----	CATGAC-TGAGTA	CTGTA--CTGTCA	CG-ATTGACATGA
E:	-----	ATG-ACAGGA-TG	-----	TGCGTTG--TGGA
F:	CAGCA-GCTG-TA	TATG-CATGATCA	CA--AAGCTGTGC	TCAATCGACA-TA
	M_1	M_2	M_3	M_4

Substitution Model



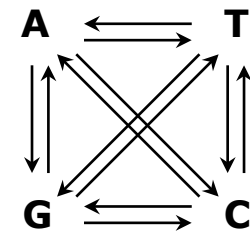
➔ PARTITION MODELS ⁽³⁾

⁽³⁾ Yang Z. 1996. Maximum-Likelihood Models for Combined Analyses of Multiple Sequence Data. J Mol Evol 42(5):587-96.

PARTITION MODELS IN MAXIMUM LIKELIHOOD (ML)

	Partition 1	Partition 2	Partition 3	Partition 4
A:	C-GTAAGCTGTAC	ATCCACA-GAGTA	GAG-ATACTGAT-	-----
B:	GAG-ATACTGTCA	-----	CAGTAAG-TGTAC	-----
C:	CTGTA--CTGCTC	-----	-----	ACACGTGAC--GA
D:	-----	CATGAC-TGAGTA	CTGTA--CTGTCA	CG-ATTGACATGA
E:	-----	ATG-ACAGGA-TG	-----	TGCGTTG--TGGA
F:	CAGCA-GCTG-TA	TATG-CATGATCA	CA--AAGCTGTGC	TCAATCGACA-TA
	M_1	M_2	M_3	M_4

Substitution Model



→ PARTITION MODELS ⁽³⁾

- Edge-Unlinked



Induced Partition Trees



Species Tree

⁽³⁾ Yang Z. 1996. Maximum-Likelihood Models for Combined Analyses of Multiple Sequence Data. J Mol Evol 42(5):587-96.

ACCOUNTING FOR PARTIAL TERRACES IN PRACTICE

- RUN TIME COMPARISON

Four programs	
RAxML	IQ-TREE
RAxML _{optU}	IQ-TREE _{PTA}

ACCOUNTING FOR PARTIAL TERRACES IN PRACTICE

- RUN TIME COMPARISON

Four programs	
RAxML	IQ-TREE
RAxML _{optU}	IQ-TREE _{PTA}

Each alignment	
10 runs	10 runs
10 runs	10 runs

ACCOUNTING FOR PARTIAL TERRACES IN PRACTICE

- RUN TIME COMPARISON

Four programs	
RAxML	IQ-TREE
RAxML _{optU}	IQ-TREE _{PTA}

Each alignment	
10 runs	10 runs
10 runs	10 runs → mean

ACCOUNTING FOR PARTIAL TERRACES IN PRACTICE

- RUN TIME COMPARISON

Four programs	
RAxML	IQ-TREE
RAxML _{optU}	IQ-TREE _{PTA}

Each alignment	
10 runs	10 runs
10 runs	10 runs → mean

Speedup	
CPU x_i / mean	CPU x_i / mean
CPU x_i / mean	CPU x_i / mean

ACCOUNTING FOR PARTIAL TERRACES IN PRACTICE

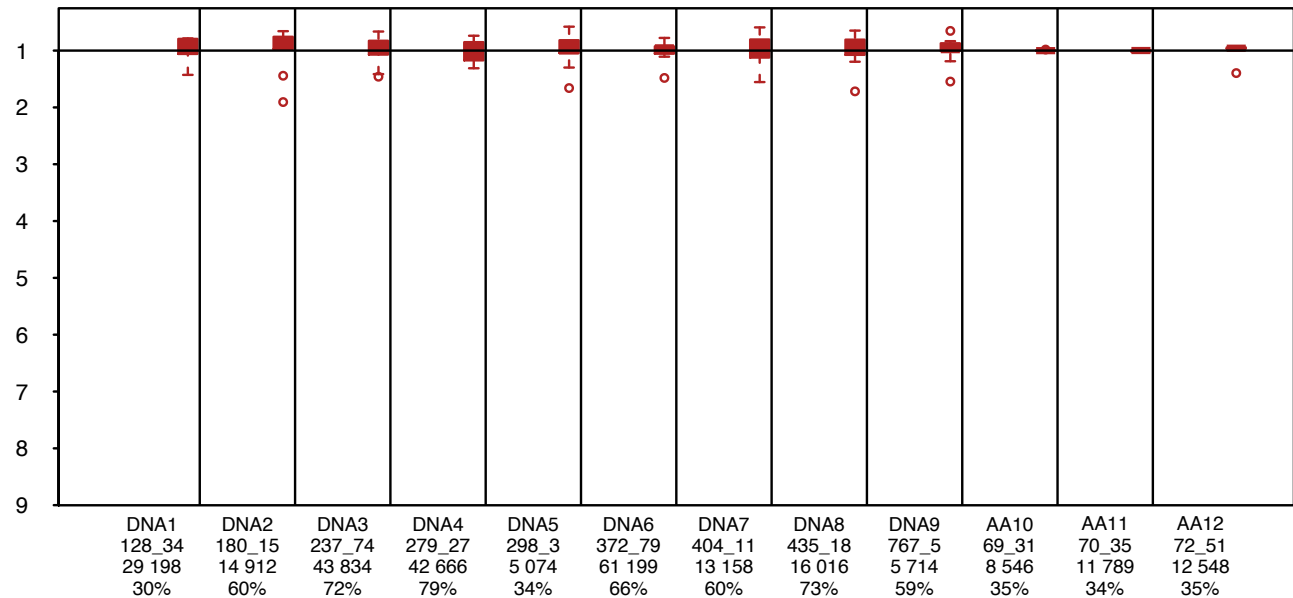
• RUN TIME COMPARISON

Edge – Unlinked

Four programs	
RAxML	IQ-TREE
RAxML _{optU}	IQ-TREE _{PTA}

Each alignment	
10 runs	10 runs
10 runs	10 runs → mean

Speedup	
CPU x_i / mean	CPU x_i / mean
CPU x_i / mean	CPU x_i / mean



IQ-TREE_{PTA}

ACCOUNTING FOR PARTIAL TERRACES IN PRACTICE

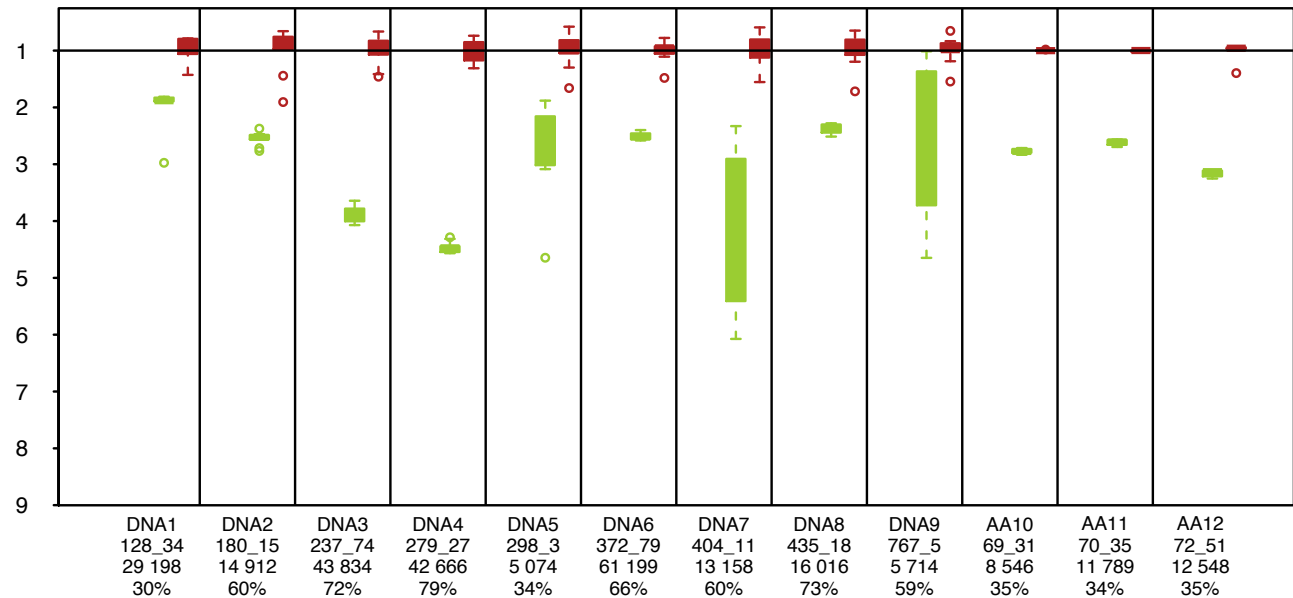
• RUN TIME COMPARISON

Edge – Unlinked

Four programs	
RAxML	IQ-TREE
RAxML _{optU}	IQ-TREE _{PTA}

Each alignment	
10 runs	10 runs
10 runs	10 runs → mean

Speedup	
CPU x_i / mean	CPU x_i / mean
CPU x_i / mean	CPU x_i / mean



IQ-TREE vs. IQ-TREE_{PTA}

ACCOUNTING FOR PARTIAL TERRACES IN PRACTICE

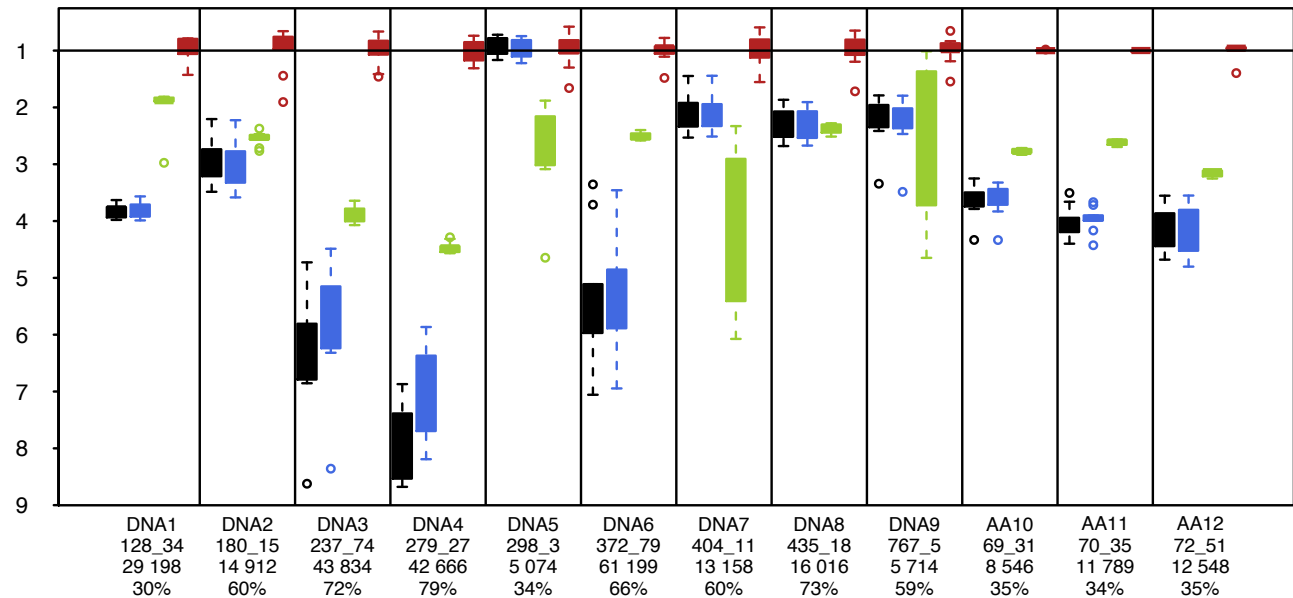
• RUN TIME COMPARISON

Edge – Unlinked

Four programs	
RAxML	IQ-TREE
RAxML _{optU}	IQ-TREE _{PTA}

Each alignment	
10 runs	10 runs
10 runs	10 runs → mean

Speedup	
CPU x_i / mean	CPU x_i / mean
CPU x_i / mean	CPU x_i / mean



SUMMARY

SUMMARY

1. **PHYLOGENETIC TERRACES**.....
an important concept that frequently occurs in practice

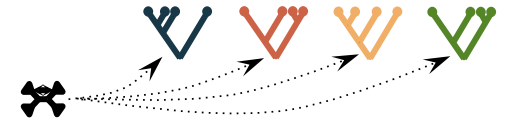


SUMMARY

1. **PHYLOGENETIC TERRACES**.....
an important concept that frequently occurs in practice



2. **DETECTION**.....
changes of induced partition trees

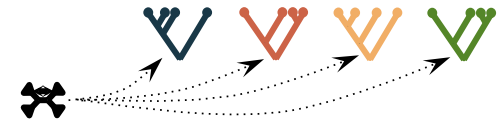


SUMMARY

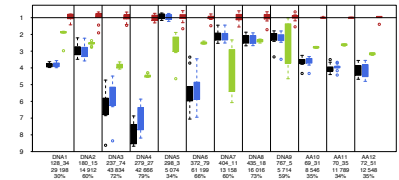
1. **PHYLOGENETIC TERRACES**.....
an important concept that frequently occurs in practice



2. **DETECTION**.....
changes of induced partition trees



3. **EFFICIENT ANALYSIS**.....
accounting for terraces and partial terraces in practice



THANK YOU FOR ATTENTION!

QUESTIONS ?



CIBIV
Center for Integrative Bioinformatics Vienna



universität
wien



FWF

Der Wissenschaftsfonds.