



**WFCC**

WORLD FEDERATION FOR CULTURE COLLECTIONS

**ICCC-12 Conference 2010**

Biological Resource Centers: gateway to biodiversity and services for innovation in biotechnology

# EVOLUTION OF THE YEAST SPECIES CONCEPT IN THE AGE OF SEQUENCING

**Marc-André Lachance**

Department of Biology  
University of Western Ontario  
London, Ontario, Canada N6A 5B7  
lachance@uwo.ca





EXPLORATION OF YEAST  
BIODIVERSITY REQUIRES A  
SOUND SPECIES CONCEPT



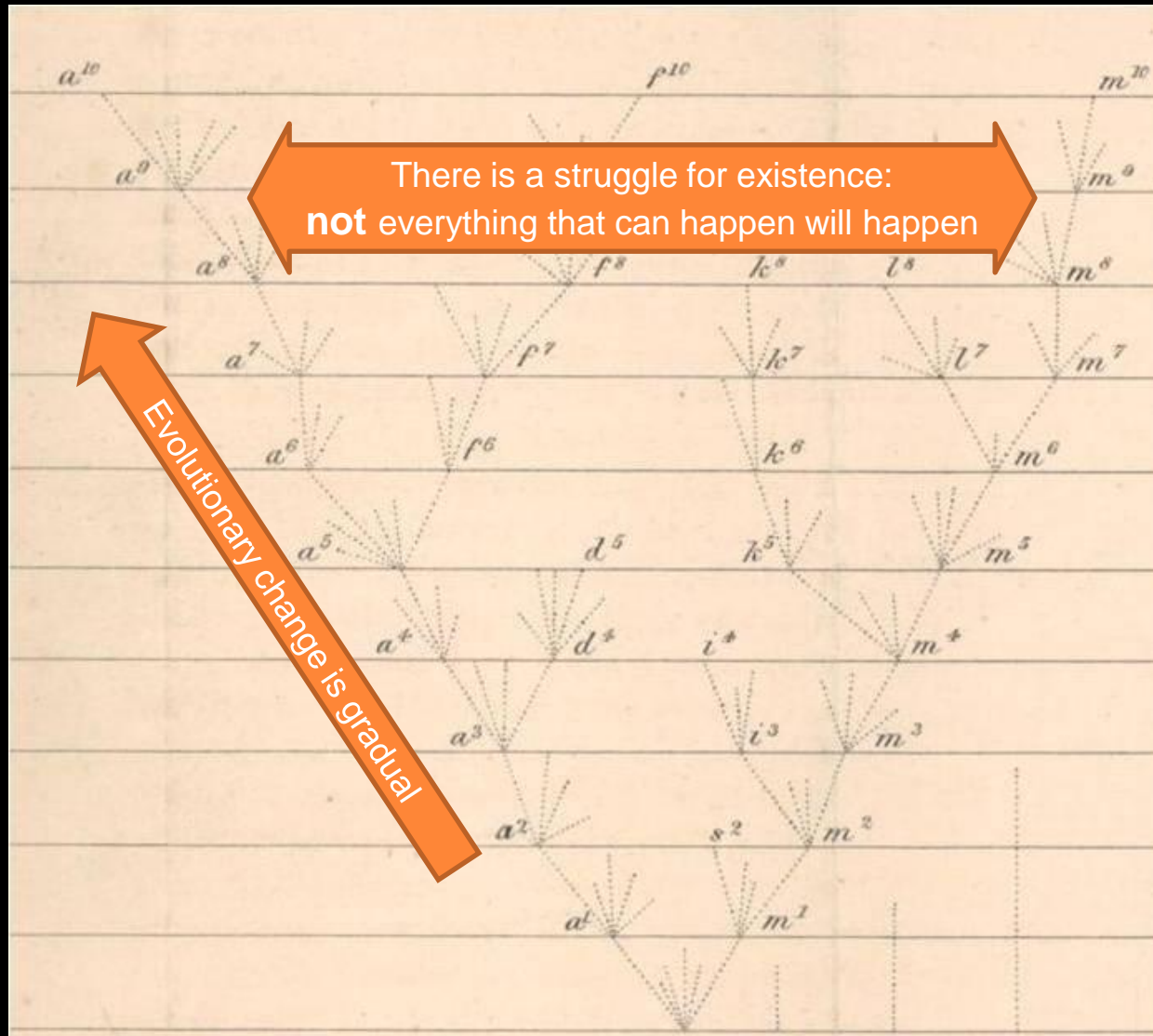
*Classis et ordo est  
sapientiae,  
species naturae opus*



*Classe e ordem são  
o trabalho da ciência,  
a espécie o da  
natureza*



# ENTERS DARWIN



# SPECIES CONCEPTS

## × Biological Species Concept

+ “Groups of interbreeding natural populations that are reproductively isolated from other such groups” (Mayr)

× Unit of evolution

× An intrinsic property of the members (objective)

× Testable/Falsifiable (*sensu* Popper)

× Limited to heterothallic, sexual species



# SPECIES CONCEPTS

## ○ Genetic Species Concept

- “Groups of individuals that share a high amount of genetic similarity and are distinct from other such groups”
  - Microbiological version uses DNA/DNA reassociation
    - 80% for yeasts
  - Calibrated on the Biological Species Concept (proxy)

○ Broadly applicable

○ Not always testable (arbitrary)

○ Onerous... sequencing is a lot easier!



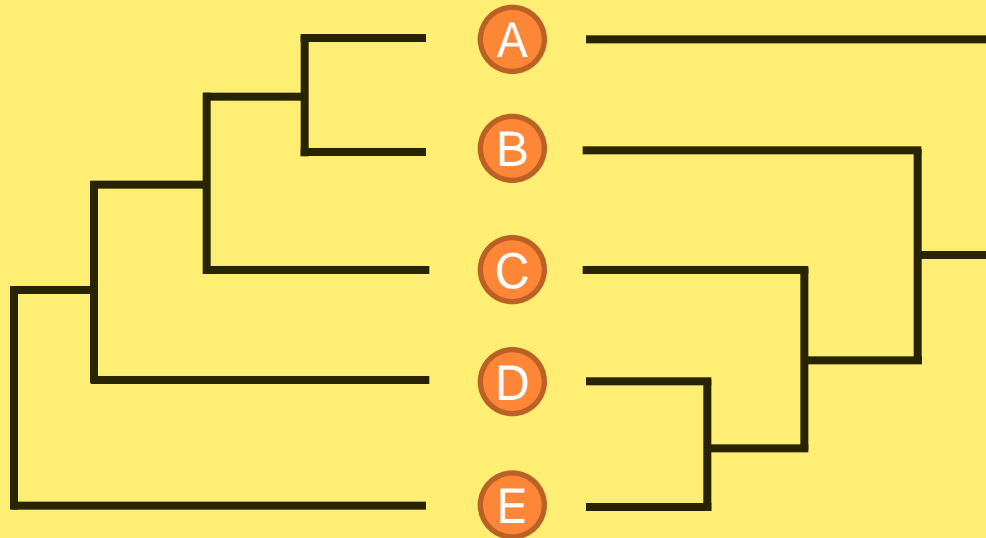
# SPECIES CONCEPTS

- Phylogenetic Species Concept
  - “The least inclusive monophyletic assemblage that can be recognized from a unique set of characteristics”
    - Can be based on DNA sequence analyses
    - Reflects the history of the members (objective)
    - Rarely testable (*sensu* Popper)
    - Inclusivity difficult to define

# EPISTEMOLOGY: PHYLOGENETIC CONCEPT

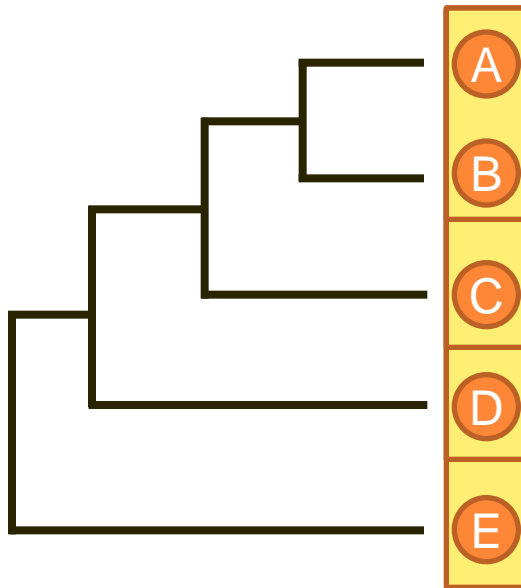
## 1. A tree is always an inference

Which is the correct tree?



# EPISTEMOLOGY: PHYLOGENETIC CONCEPT

2. Assuming that this is the correct tree, which is the least inclusive assemblage?



The amount of differences considered necessary to give to any two forms the rank of species cannot be defined.

Darwin 1859



# SPECIES CONCEPTS

## × *Sequence Divergence* Species Concept

+ “Groups of individuals that differ by 3 or fewer substitutions in the D1/D2 LSU rRNA gene and differ from other such groups by 1% or more divergence

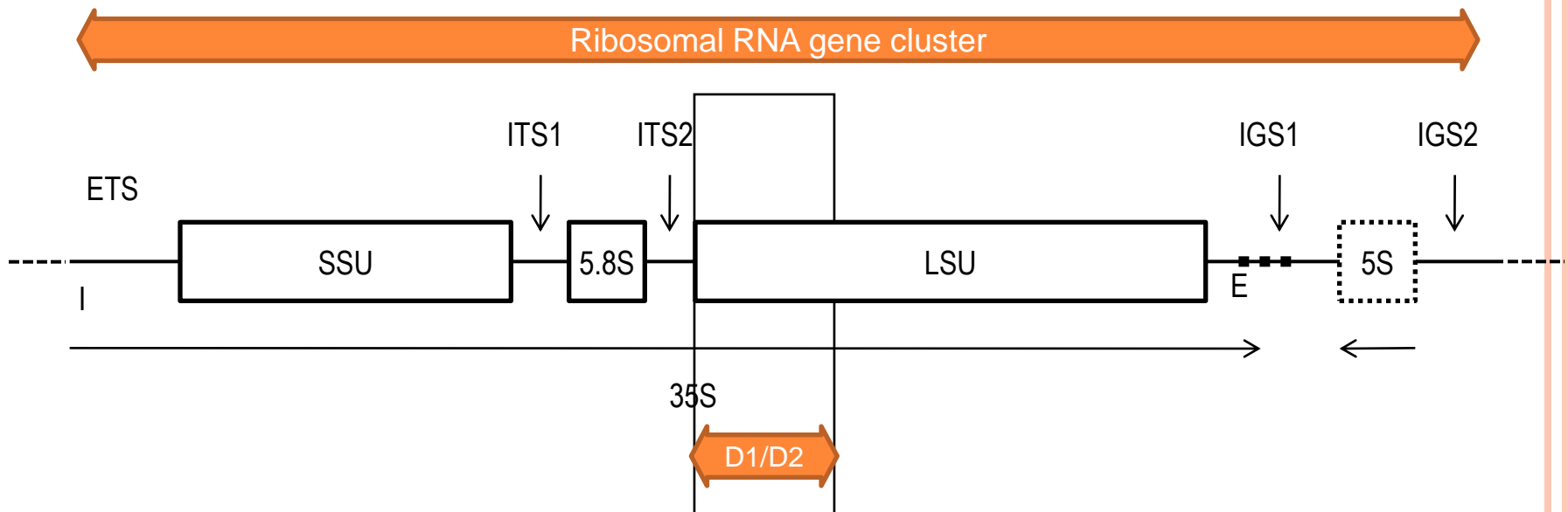
× Can be based on DNA sequence analyses

× Broadly applicable

× Some notorious exceptions

× Not always testable (arbitrary)

# YEAST BARCODING SEQUENCE



# SPECIES CONCEPTS

## × *Sequence Divergence* Species Concept

+ “Groups of individuals that differ by 3 or fewer substitutions in the D1/D2 LSU rRNA gene and differ from other such groups by 1% or more divergence

× Can be based on DNA sequence analyses

× Broadly applicable

\* Some notorious exceptions



× Not always testable (arbitrary)



## SEQUENCE DIVERGENCE CONCEPT - EXCEPTIONS

- *Metschnikowia agaves*

- Holotype and allotype differ by 5 substitutions
  - Kurtzman & Robnett 1998

- *Clavispora lusitaniae*

- Holotype *et al* differ by > 30 substitutions from allotype *et al*
- Some strains contain both D2 variants
  - Lachance et al. 2003



# SPECIES CONCEPTS

## × *Sequence Divergence* Species Concept

+ “Groups of individuals that differ by 3 or fewer substitutions in the D1/D2 LSU rRNA gene and differ from other such groups by 1% or more divergence

× Can be based on DNA sequence analyses

× Broadly applicable

\* Some notorious exceptions

× Not always testable (arbitrary) ←

## SEQUENCE DIVERGENCE CONCEPT - EPISTEMOLOGY

### ○ Induction

- “Taking into account the variation seen in the preceding comparisons, it is **predicted** that strains showing greater than 1% substitutions in the ca. 600-nucleotide D1/D2 domain **are likely to be different species** and that strains with 0–3 nucleotide differences are either conspecific or sister species.” (Kurtzman and Robnett 1998)

### ○ Generalization

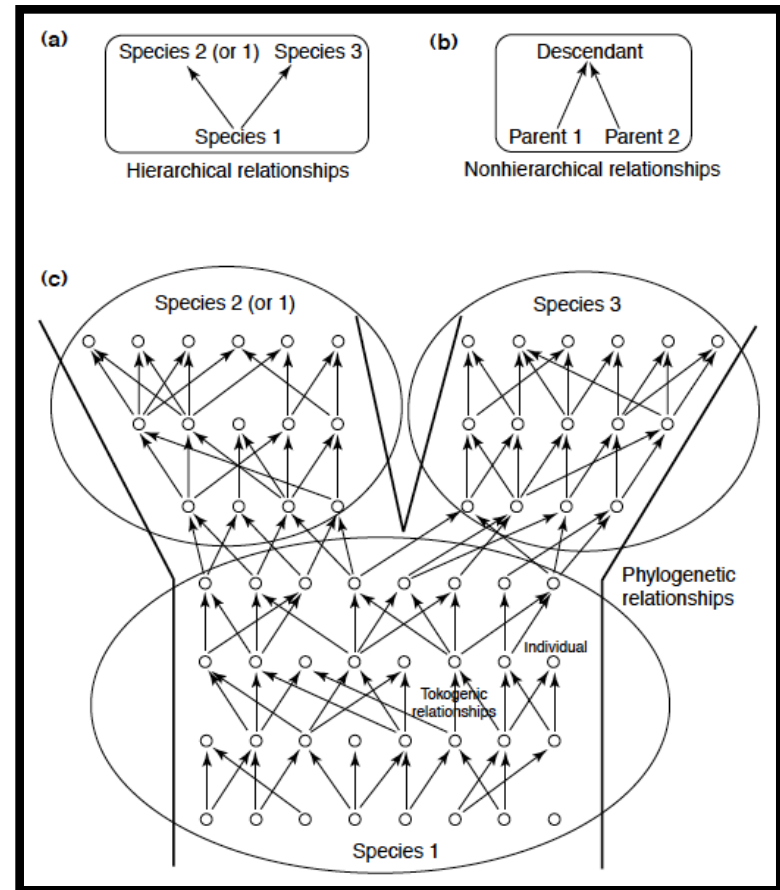
- “Kurtzman and Robnett (1998) **demonstrated** for ascomycetous yeasts that strains differing by more than 1% substitutions in the D1/D2 domain **represent separate species.**” (Kurtzman and Droby 2001)



# TOKOGENY VS PHYLOGENY

“The only distinction between species and well-marked **varieties** is, that the latter **are** known, or believed, to be **connected at the present day** by intermediate gradations, whereas **species were formerly thus connected.**”

Darwin 1859



Posada & Crandall 2001

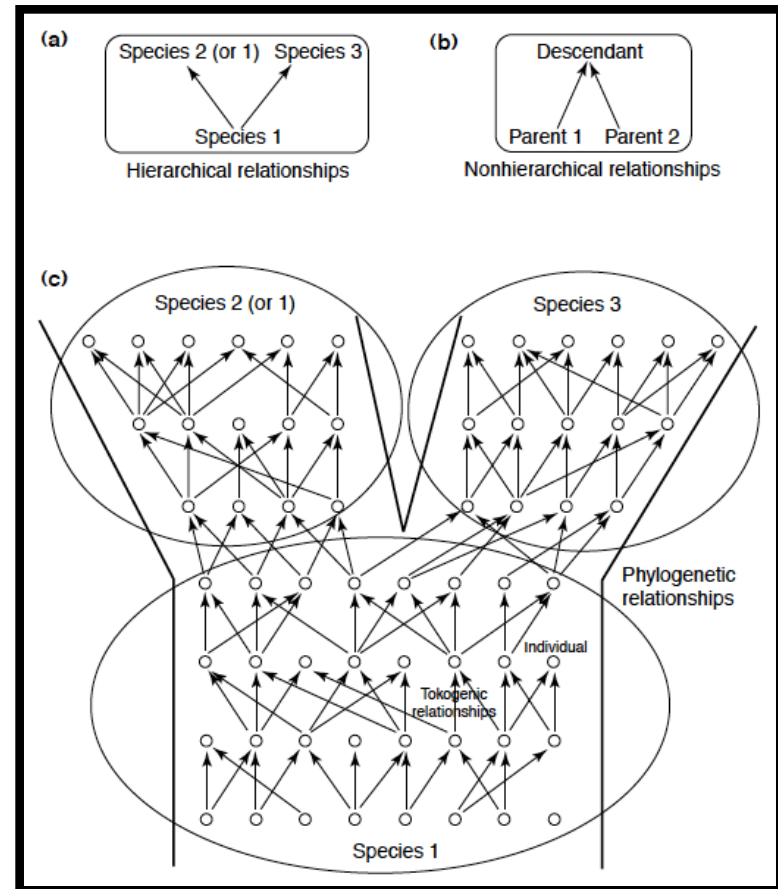
ex Hennig 1966

# TOKOGENY VS PHYLOGENY

Varieties are  
connected *at  
the present day*

species were  
*formerly* thus  
connected

Darwin 1859

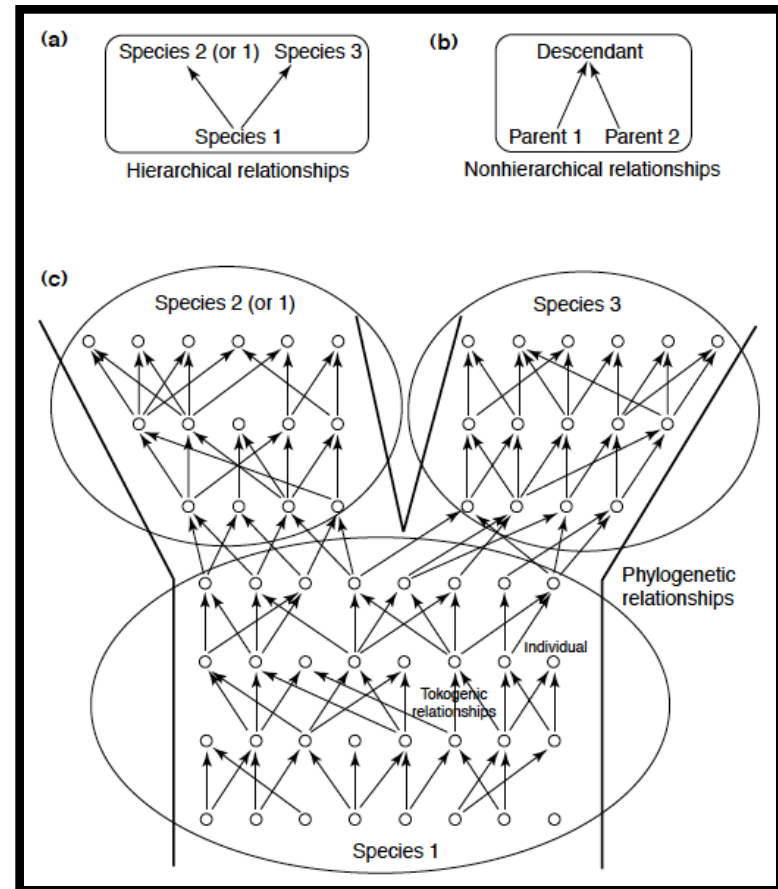


Posada & Crandall 2001

ex Hennig 1966

# TOKOGENY VS PHYLOGENY

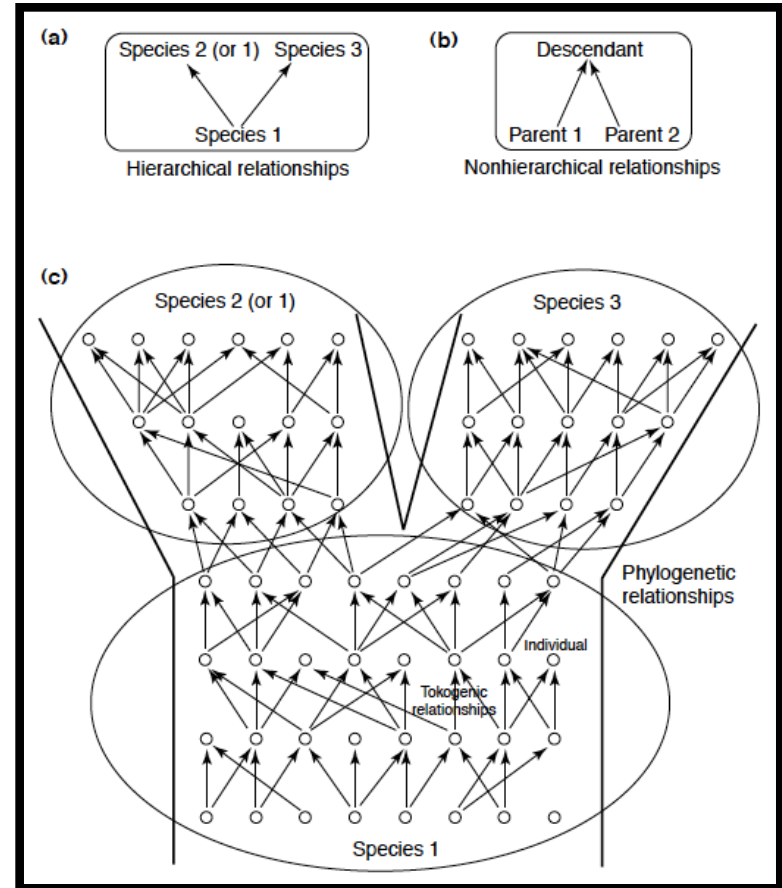
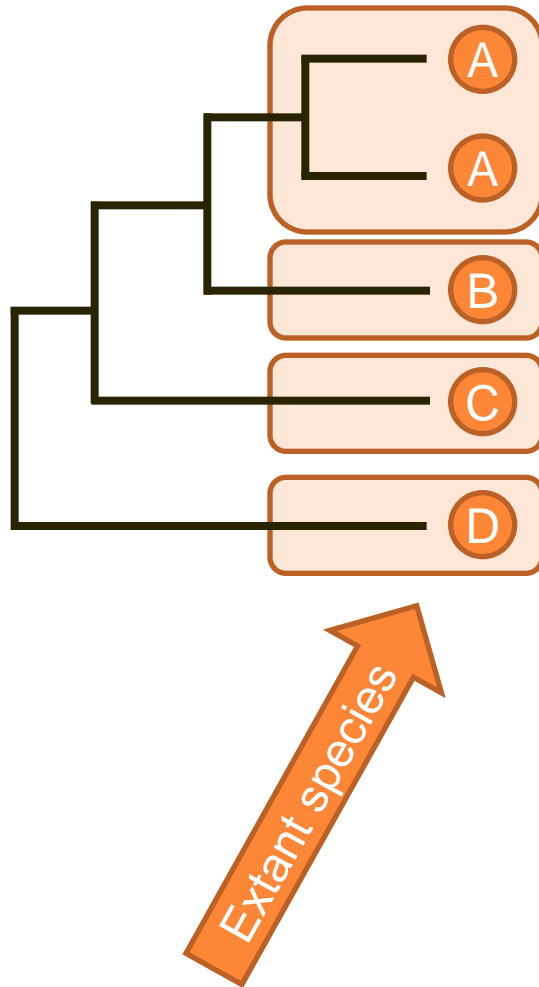
- ✗ Sequence relationships *between* species are phylogenetic, hierarchical, *i.e.*, **tree-like**
- ✗ Sequence relationships *within* species are tokogenetic, nonhierarchical, *i.e.*, **network-like**



Posada & Crandall 2001

ex Hennig 1966

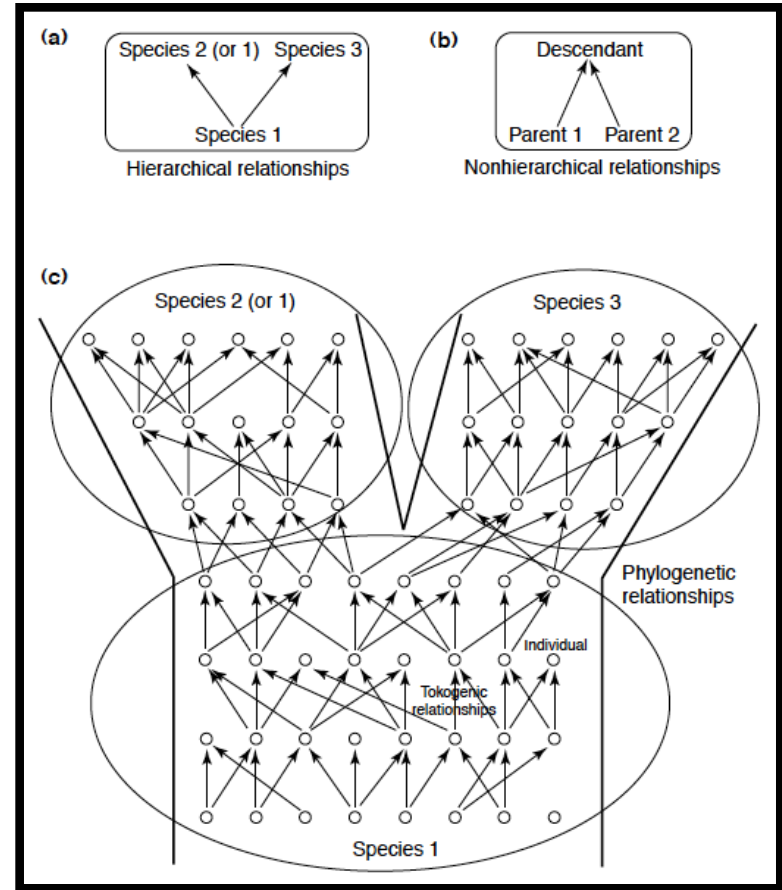
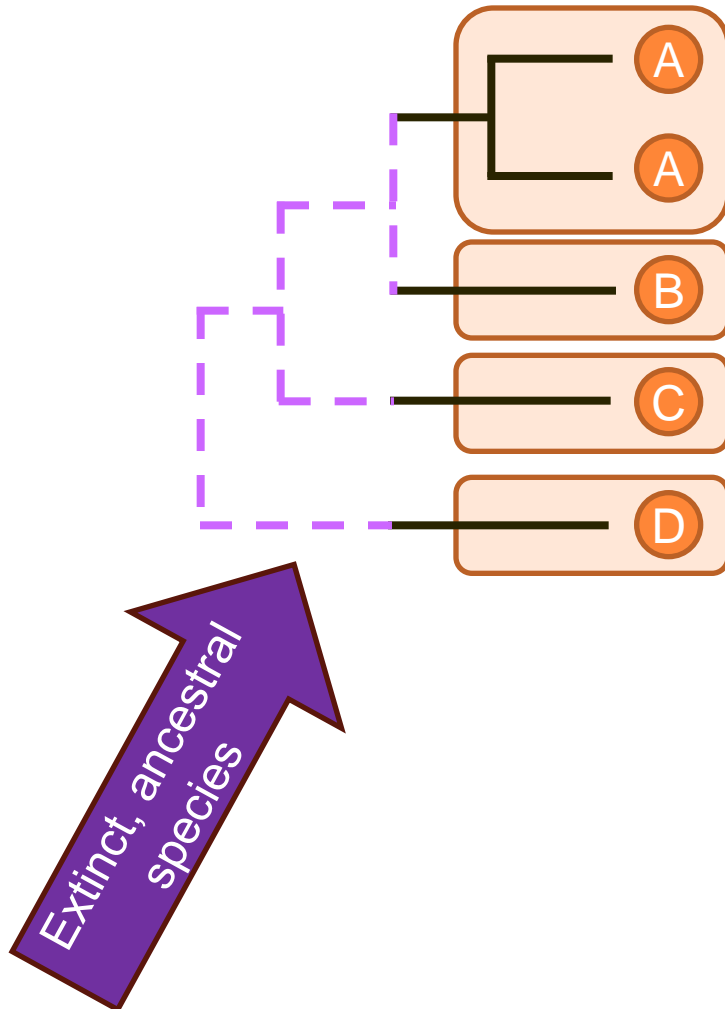
# TOKOGENY VS PHYLOGENY



Posada & Crandall 2001

ex Hennig 1966

# TOKOGENY VS PHYLOGENY

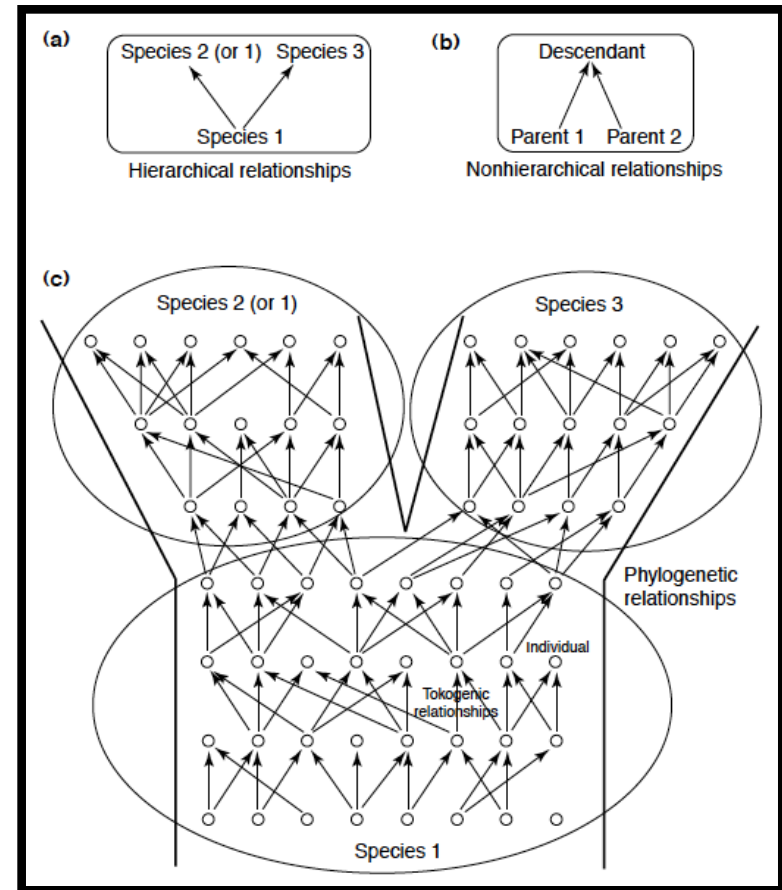
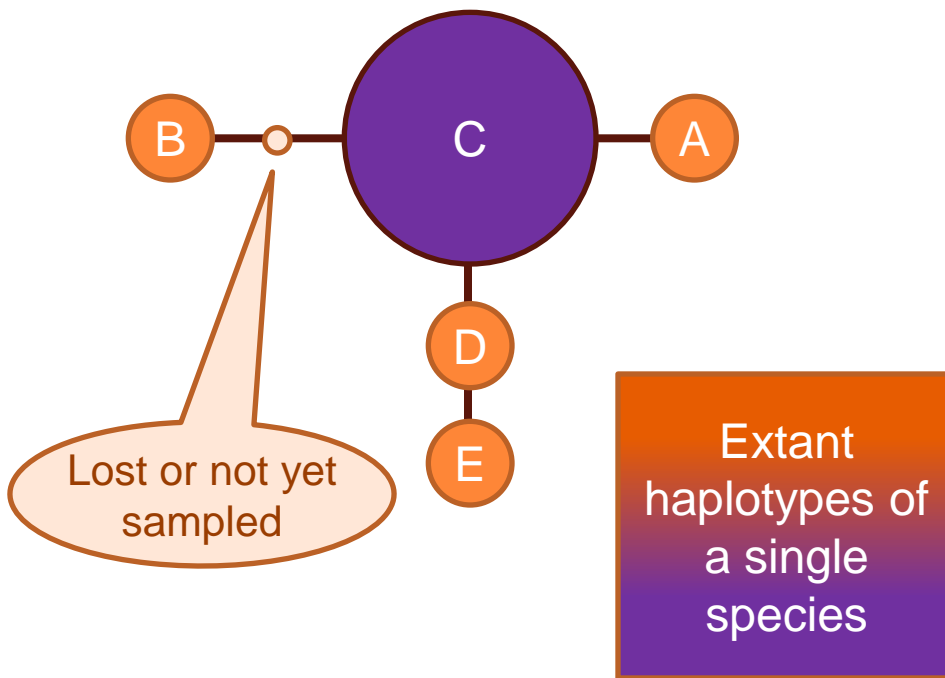


Ancestral sequences no longer exist!

Posada & Crandall 2001

ex Hennig 1966

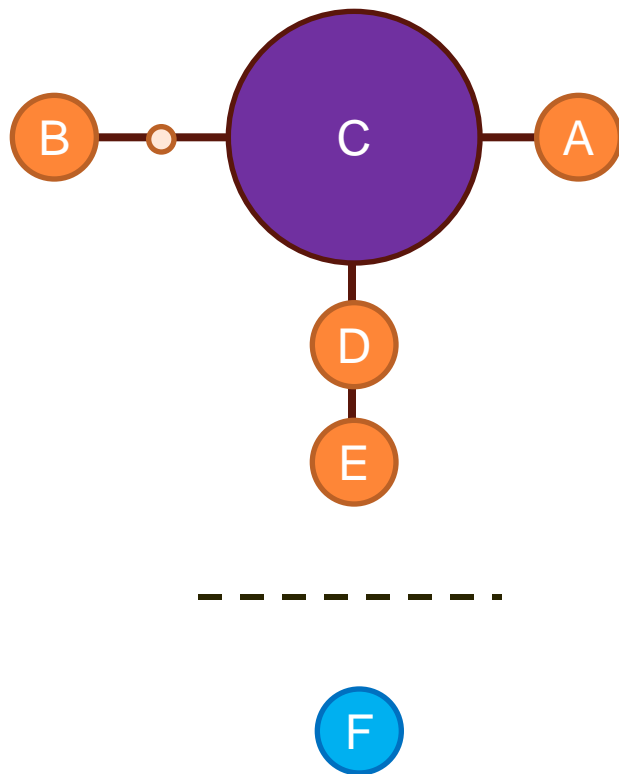
# TOKOGENY VS PHYLOGENY



Posada & Crandall 2001

ex Hennig 1966

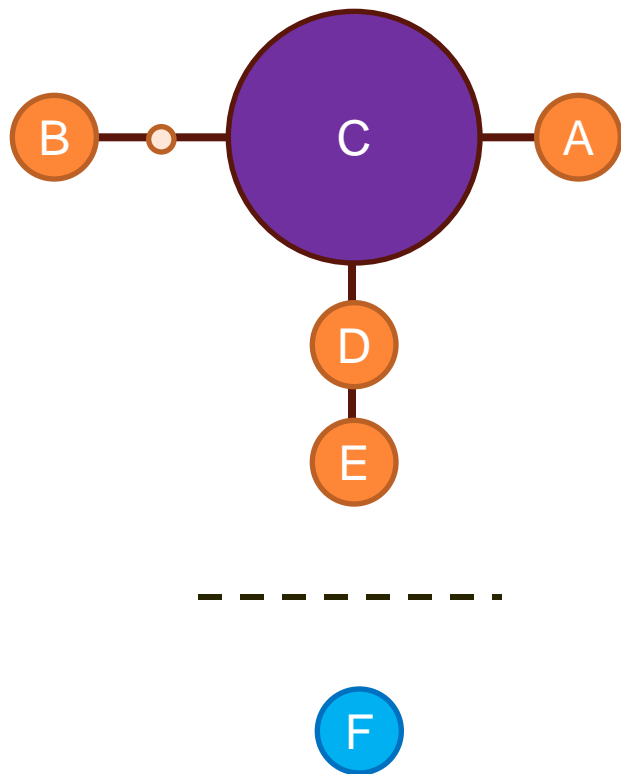
# TOKOGENY VS PHYLOGENY



The program **TCS** is an implementation of **T**empleton, **C**randall, and **S**ing's (1992) haplotype network analysis with a statistical parsimony test for membership

Sequences are accepted in a network if the probability that each step represents a single substitution is greater than a specified  $P$  (0.95)

# TOKOGENY VS PHYLOGENY



biology  
**letters**

Evolutionary biology

*Biol. Lett.* (2007) 3, 509–512

doi:10.1098/rsbl.2007.0307

Published online 24 July 2007

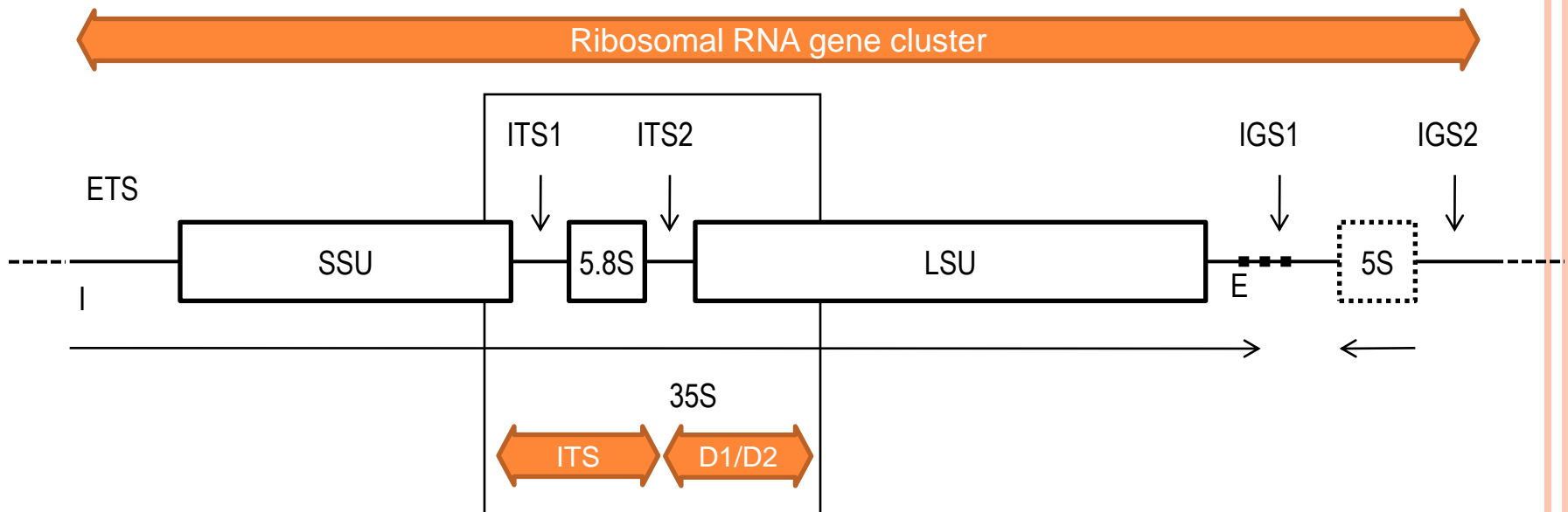
## Things fall apart: biological species form unconnected parsimony networks

Michael W. Hart\* and Jennifer Sunday

*Department of Biological Sciences, Simon Fraser University*

78% of species studied identified correctly by TCS ( $N = 663$  across Eukaryota)

# YEAST BARCODING SEQUENCE

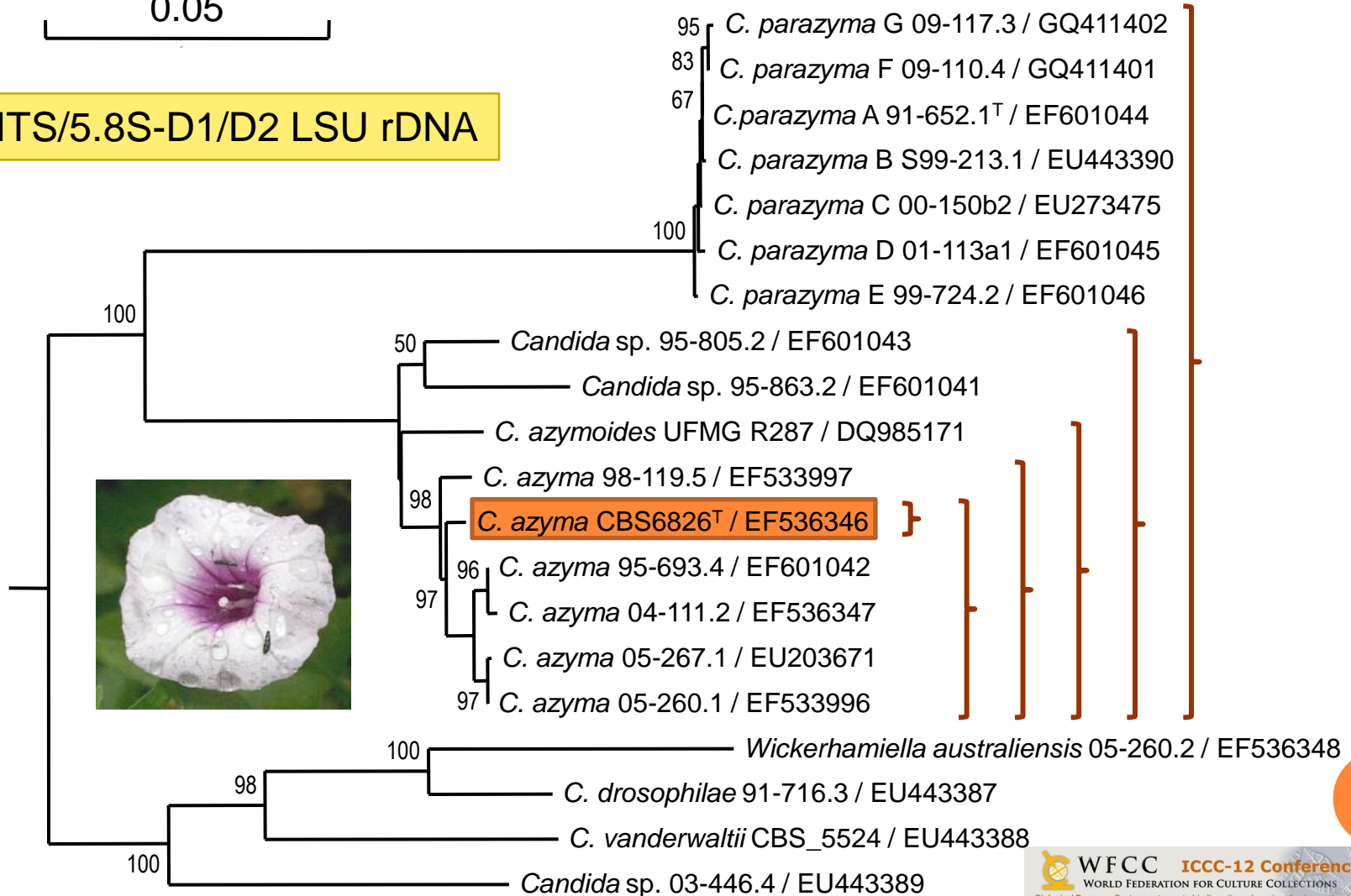


# CASE: *CANDIDA AZYMA*

Lachance,  
Dobson,  
D Wijayanayaka,  
Smith  
2010

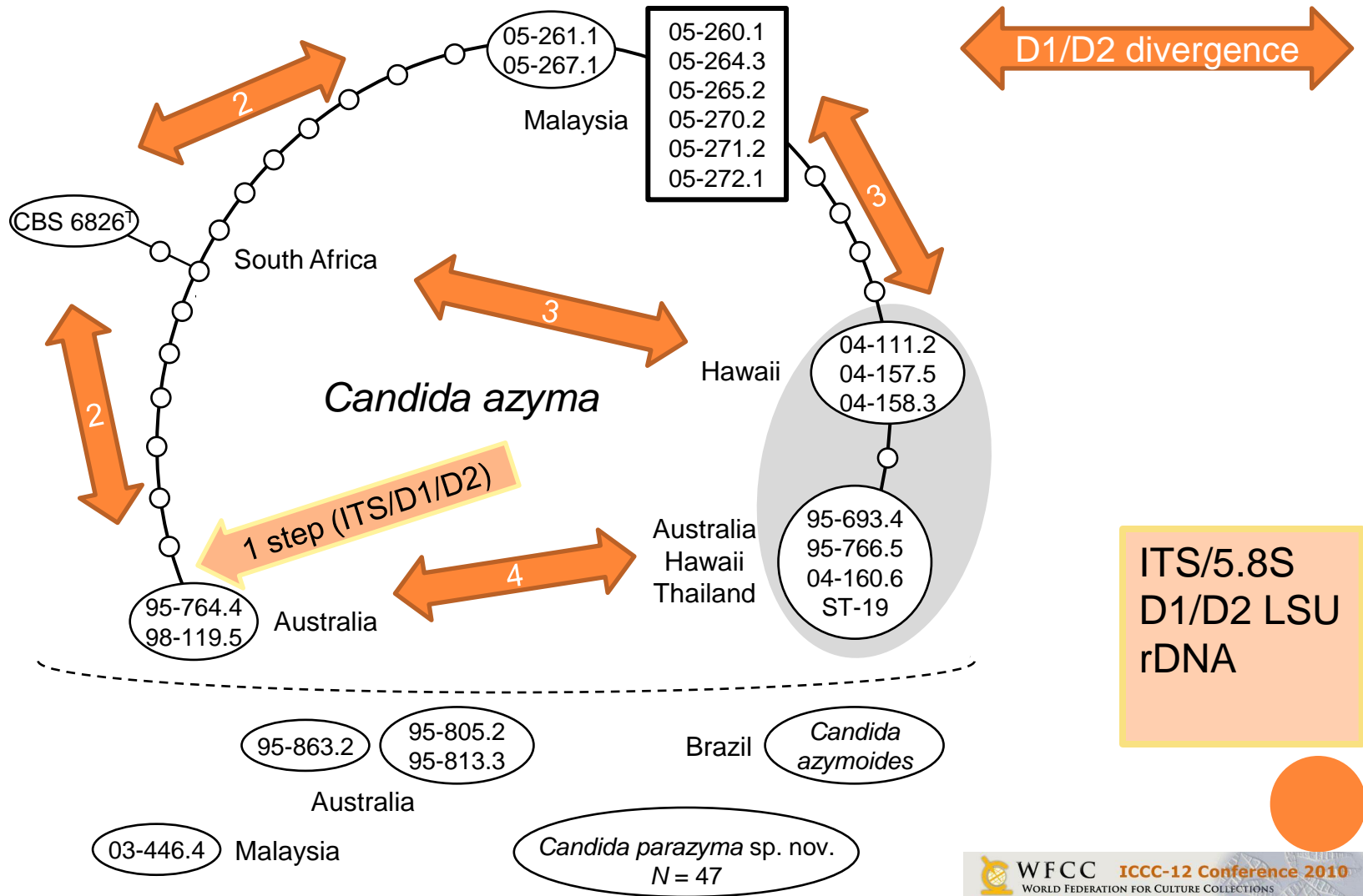
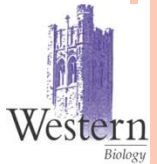
0.05

ITS/5.8S-D1/D2 LSU rDNA



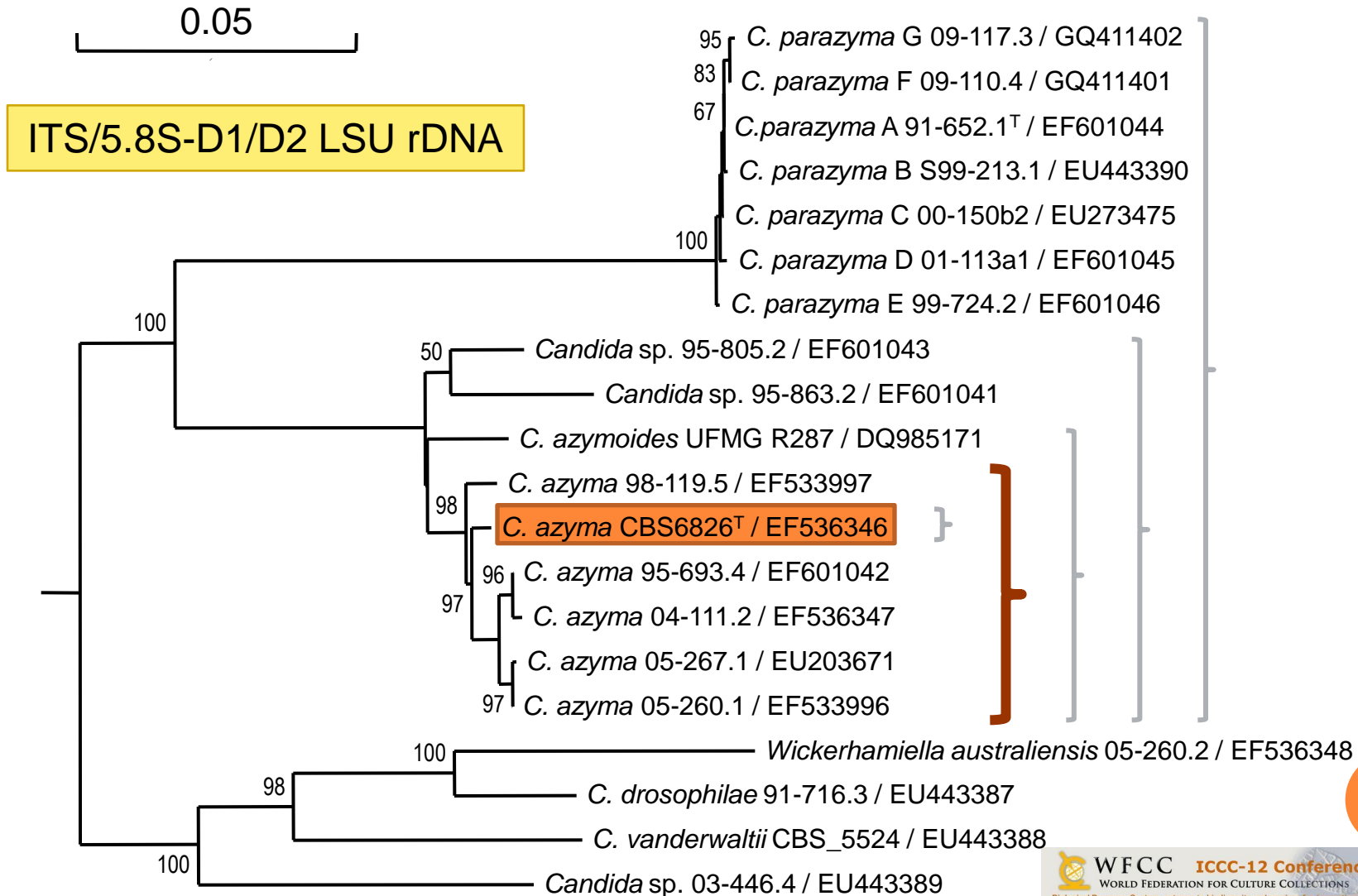
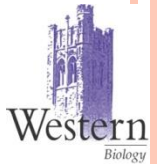
# CASE: *CANDIDA AZYMA*

Lachance,  
Dobson,  
D Wijayanayaka,  
Smith  
2010



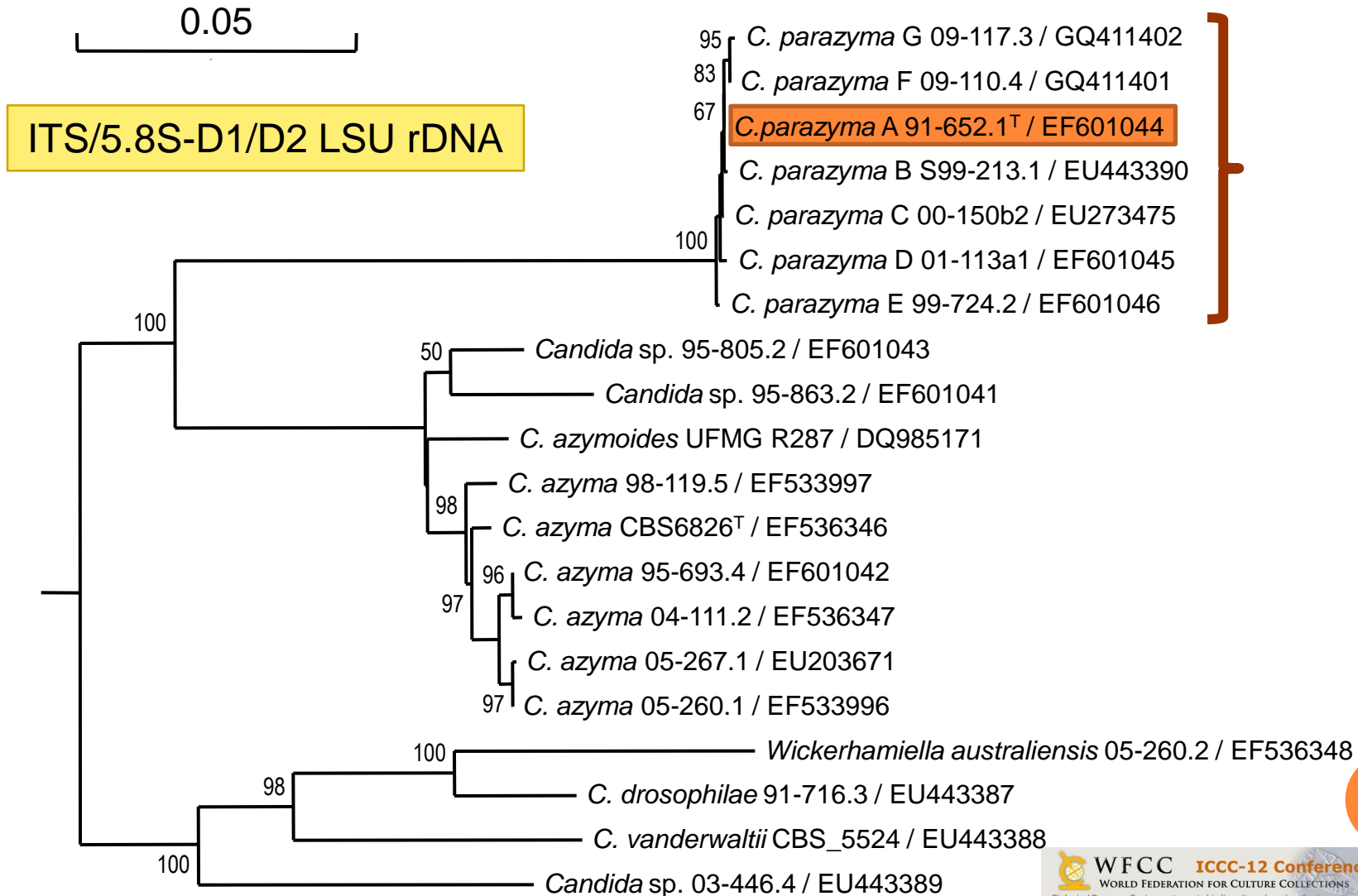
# CASE: *CANDIDA AZYMA*

Lachance,  
Dobson,  
D Wijayanayaka,  
Smith  
2010



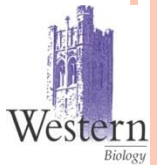
# CASE: *CANDIDA PARAZYMA*

Lachance,  
Dobson,  
D Wijayanayaka,  
Smith  
2010

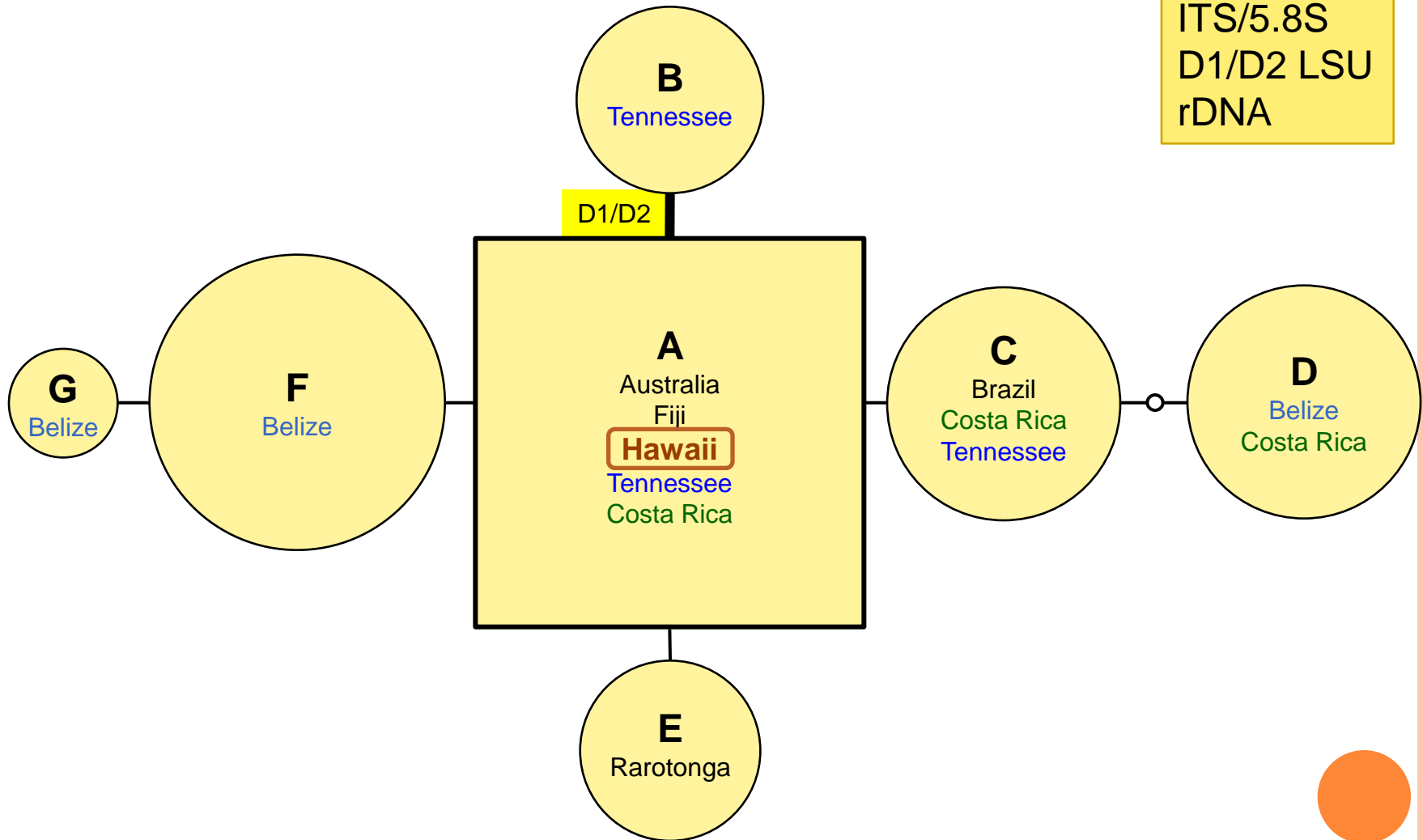


# CASE: *CANDIDA PARAZYMA*

Lachance,  
Dobson,  
D Wijayanayaka,  
Smith  
2010

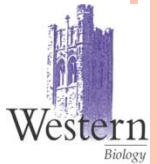


ITS/5.8S  
D1/D2 LSU  
rDNA

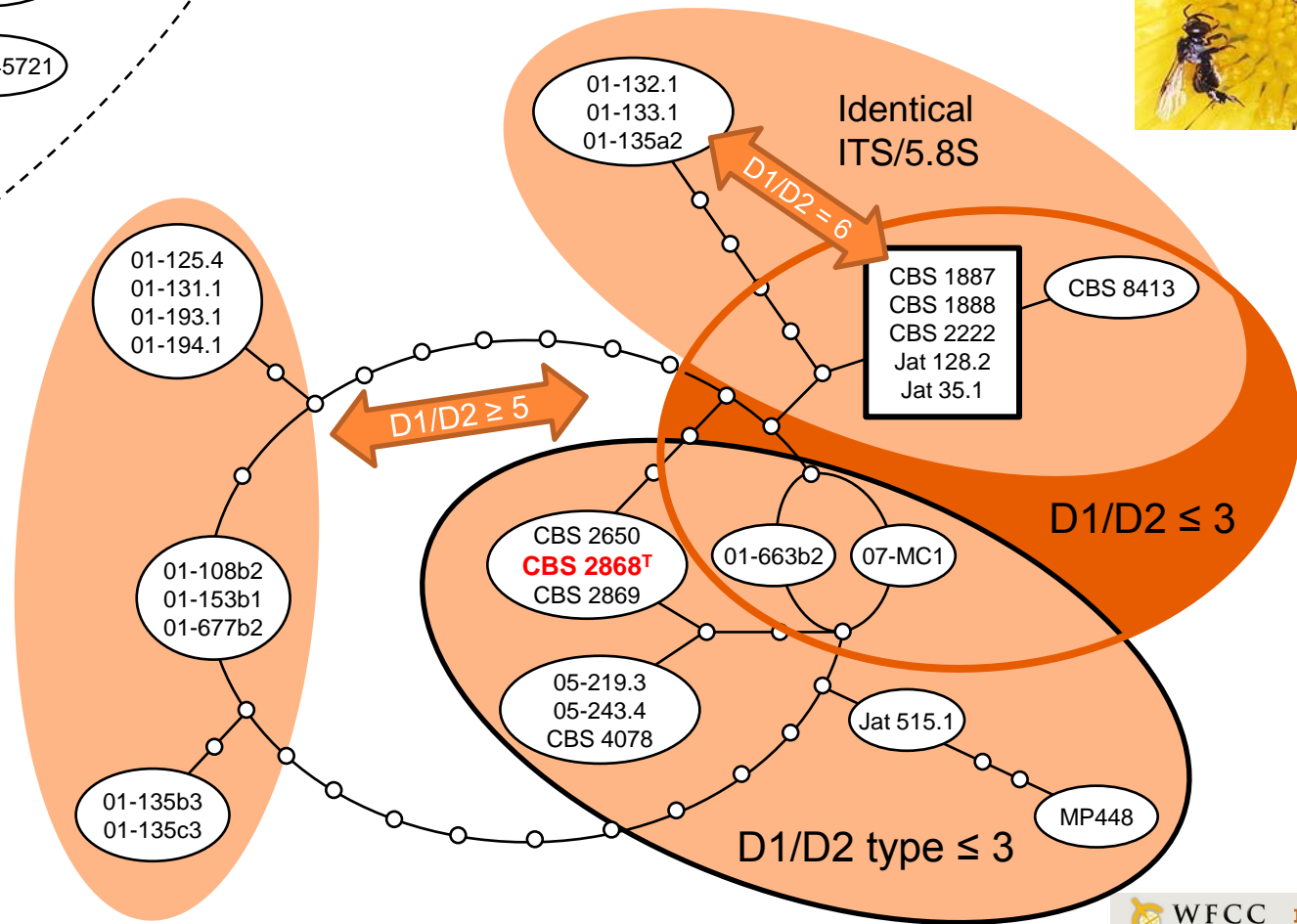


# CASE: *CANDIDA APICOLA*

Lachance,  
Dobson,  
D Wijayanayaka,  
Smith  
2010



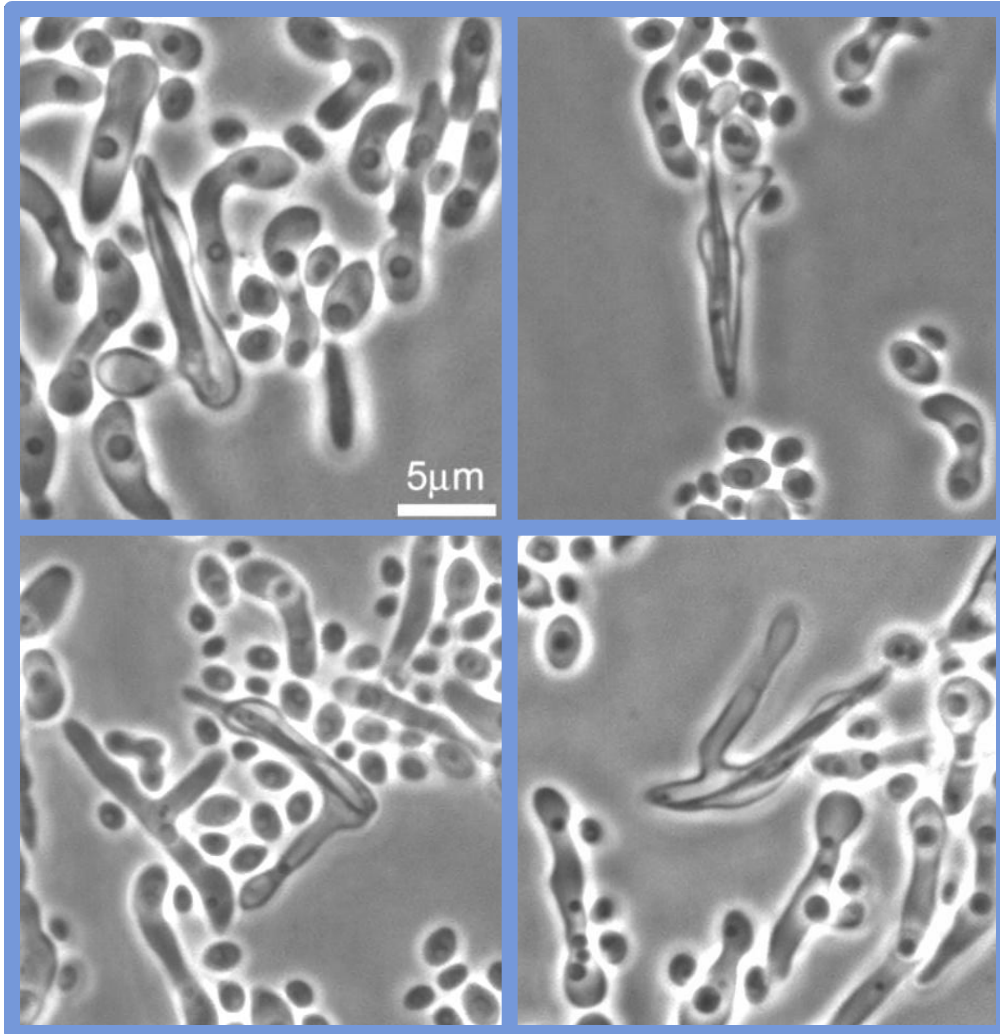
CBS 4353  
MUCL 45721



ITS/5.8S  
D1/D2 LSU  
rDNA

# CASE: *METSCHNIKOWIA AGAVES*

Bundus & Lachance  
unpublished

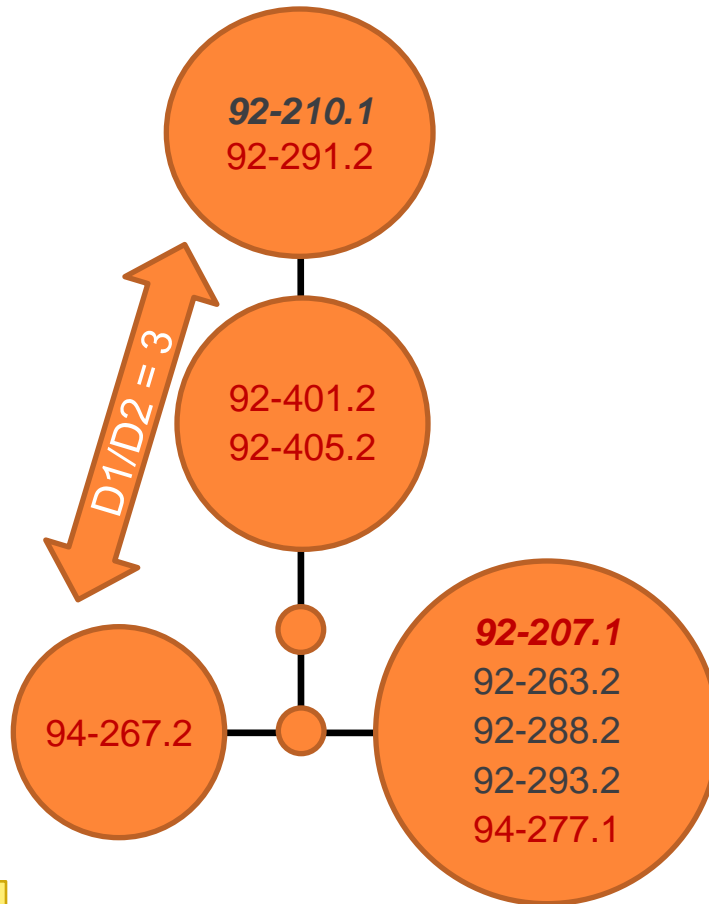


$N = 10$

All (6×4) compatible crosses  
produce asci with two  
ascospores

# CASE: *M. AGAVES*

Bundus & Lachance  
unpublished



## Extrapolations:

**Predicted** maximum number of  
D1/D2 variants in species  $\approx 6$

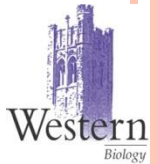
**Predicted** maximum number of  
D1/D2 substitutions in species  $\approx 4-5$



ITS/5.8S  
D1/D2 LSU  
rDNA

# CASE: STARMERELLA BOMBICOLA

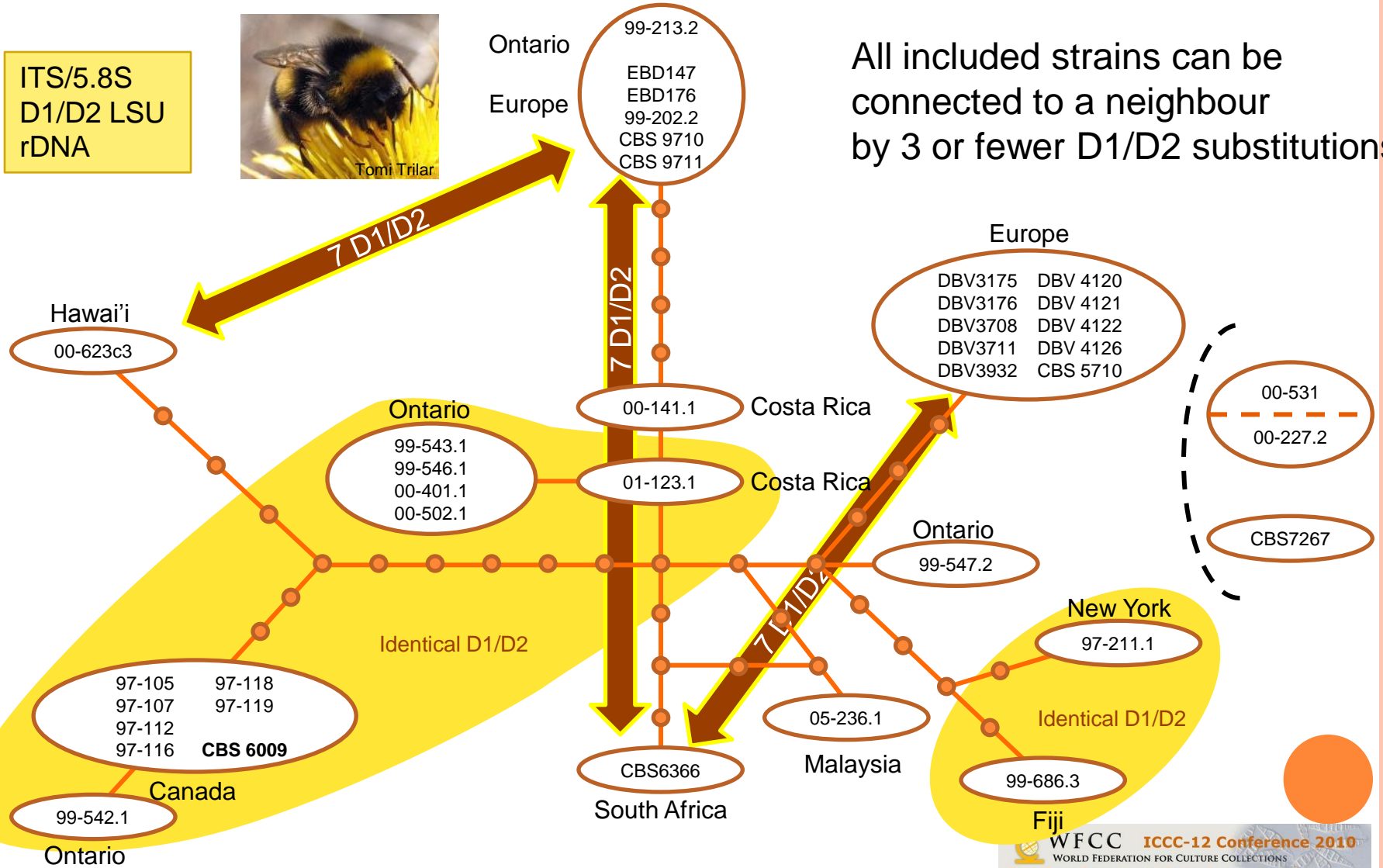
T. Wijayanayaka  
& Lachance  
unpublished



ITS/5.8S  
D1/D2 LSU  
rDNA

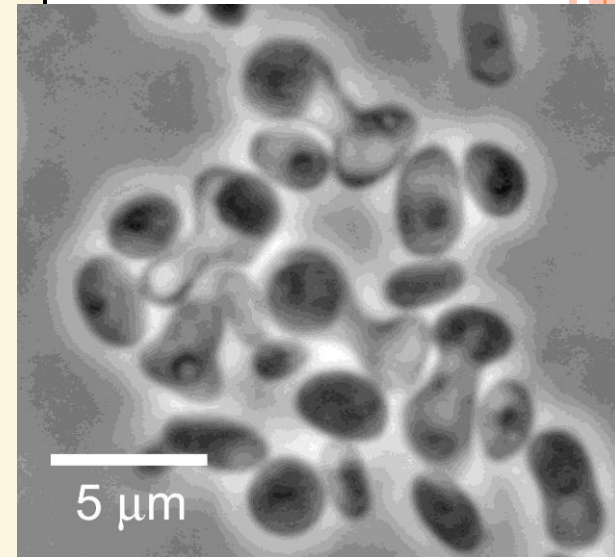
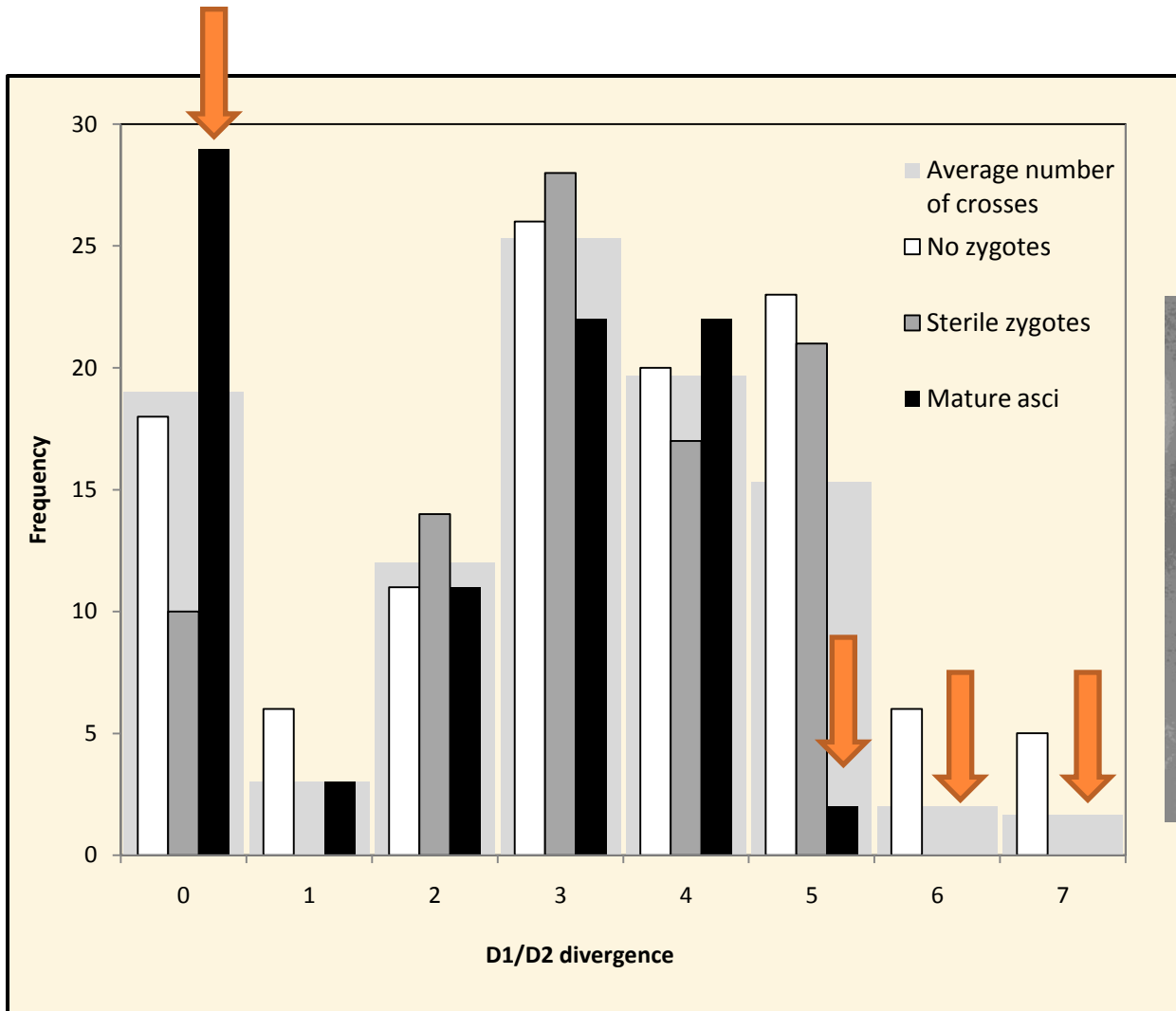
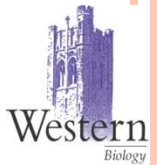


All included strains can be connected to a neighbour by 3 or fewer D1/D2 substitutions



# CASE: *STARMERELLA BOMBICOLA*

T. Wijayanayaka  
& Lachance  
unpublished



# SUMMARY

- Barcode DNA sequencing
  - has become the standard for yeast identification
  - provides preliminary phylogenetic information
- Application to species delineation must be rooted in a sound theoretical framework
  - Adequate sampling of species is essential
  - ITS/5.8S-D1/D2 haplotype networks help in discriminating between tokogenetic (within-species) and phylogenetic (between species) relationships



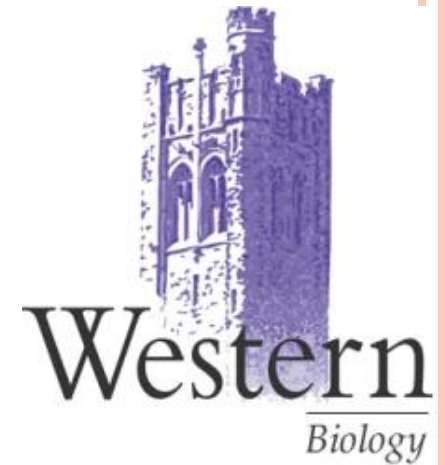
# ACKNOWLEDGEMENTS

## ○ Lab work

- Jessica Dobson
- Dilini Wijayanayaka
- Allison Smith
- Joanna Bundus
- Tishara Wijayanayaka

## ○ Inspiration

- Daria Koscinski
- Clete Kurtzman
- Carlos Rosa
- Tom Starmer



**NSERC  
CRSNG**



**WFCC**

**ICCC-12 Conference 2010**

WORLD FEDERATION FOR CULTURE COLLECTIONS

Biological Resource Centers: gateway to biodiversity and services for innovation in biotechnology