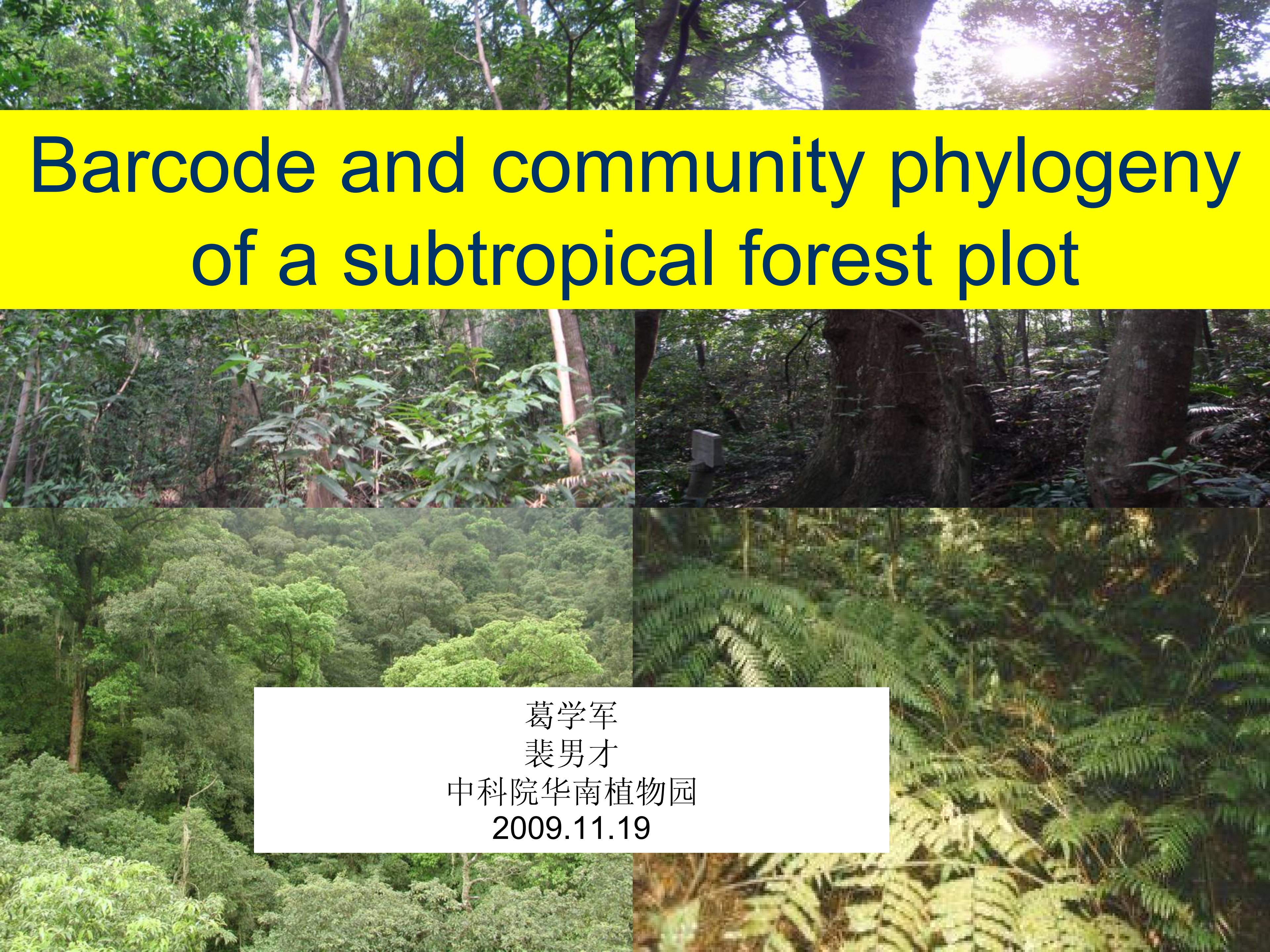


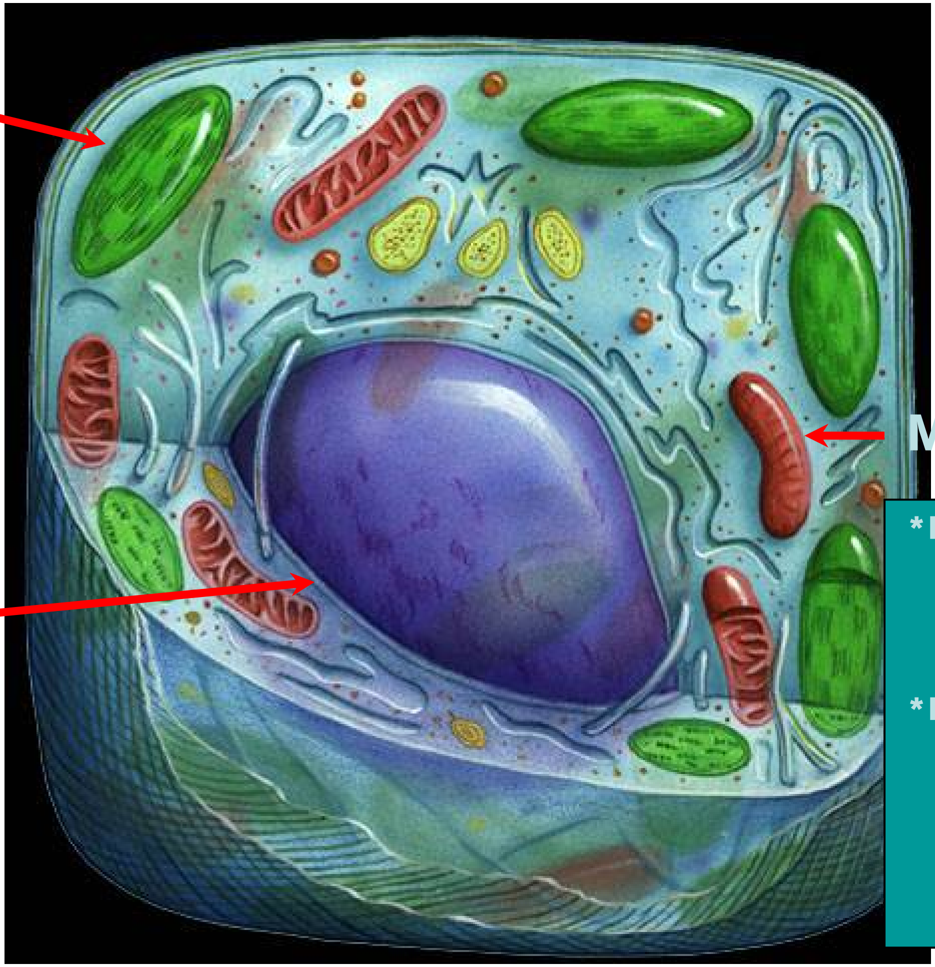
Barcode and community phylogeny of a subtropical forest plot

葛学军
裴男才
中科院华南植物园
2009.11.19

DNA Barcodes

A short universal gene sequence taken from a standardized portion of the genome used to identify species

Three Genomes of Plant Cells for Barcode Candidates



Chloroplast

- *High copy number
- *Conserved structure
- *Diversity of substitution rates across genes, introns, and intergenic spacers

Nuclear

- *Contain the most variable loci
- *Problems with multi-gene families
- *Single-copy genes often technically difficult

Mitochondrial

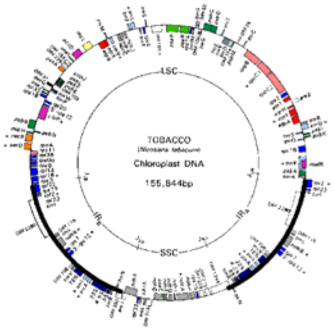
- *Locus of choice for animal barcoding is mitochondrial COI
- *Limitations with plants
  - -Low divergence
  - -Rapid genome rearrange-ments

Challenges for plant barcoding

- Choosing a plant barcode
  - Plant mtDNA substitution rates much slower than in animals (*CO1* does not work well in plants)
  - cpDNA a better alternative, but still slower than animal mtDNA

- Discriminatory power
  - Plant species frequently hybridise, and there are many examples of recent speciation

Choosing the plant barcoding region



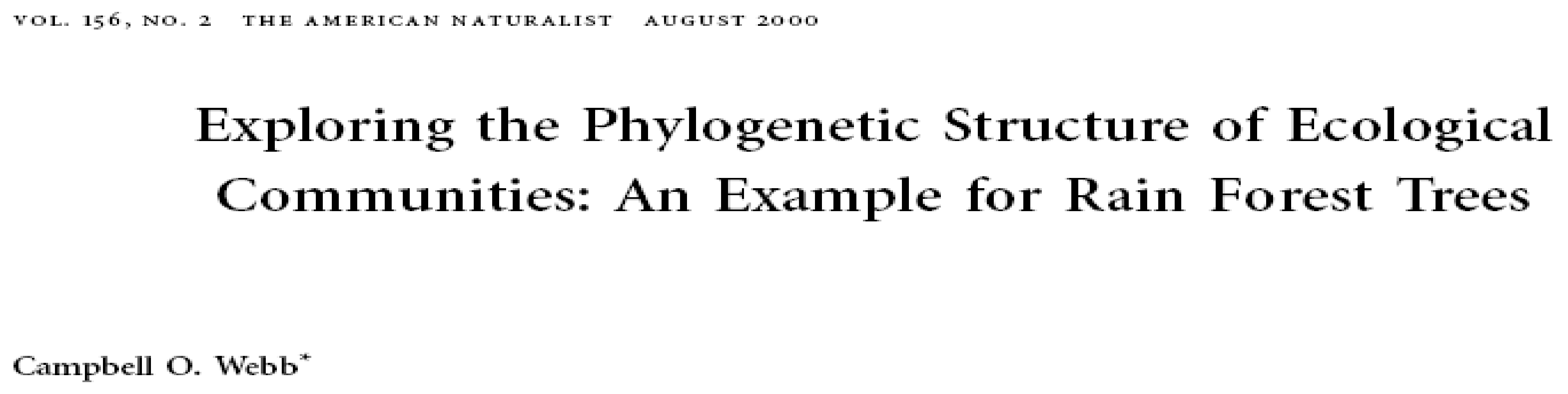
Proposed solutions:

Chase et al (2007): (Taxon 56, 295-299)	*rpoC1 + rpoB + matK* *rpoC1 + matK + trnH-psbA*
Kim et al (2007): (unpublished)	*matK + atpF-H + psbK-I* *matK + atpF-H + trnH-psbA*
Kress/Erickson (2007): (PLoS ONE 2, e508)	*rbcL + trnH-psbA*
Lahaye *et al.* (2008): (PNAS 105, 2923–2928 *Nature Precedings http://hdl.handle.net/10101/npre.2008.1896.1*)	*matK* (or *matK + trnH-psbA*)
Plant working group (2007) (minutes of Taipei meeting)	*matK + trnH-psbA + atpF-H*
Plant working group (2009) **(Mexico meeting)**	***rbcL + matK (+ trnH-psbA, ITS)***

Applications of plant barcoding

- Species identification and Cryptic species discovery
- Identification of plant material in diets (animal, insect)
- Identification and origin study of invasive species
- Root system study
- CTFS plots

Phylogeny and Community Ecology



VOL. 156, NO. 2 THE AMERICAN NATURALIST AUGUST 2000

Exploring the Phylogenetic Structure of Ecological Communities: An Example for Rain Forest Trees

Campbell O. Webb*

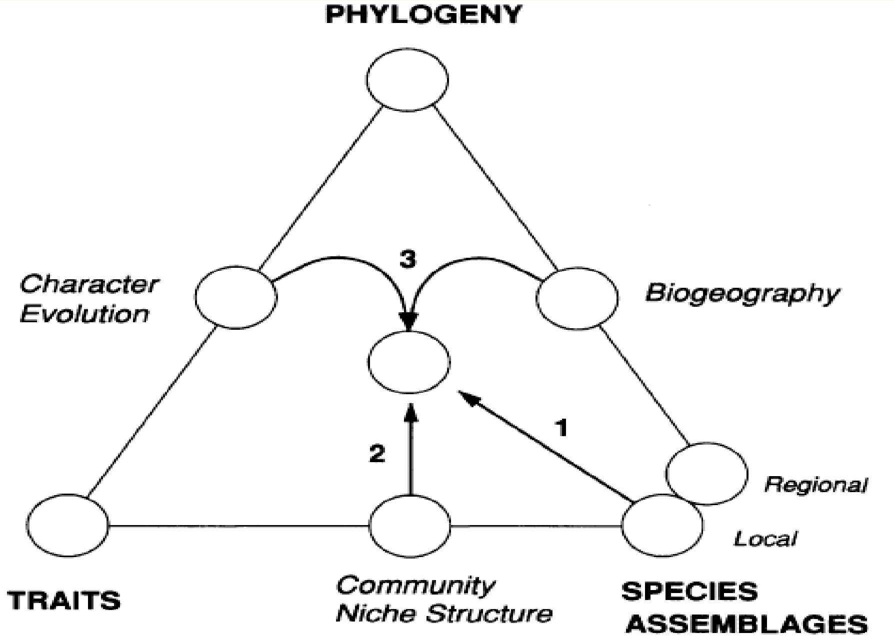
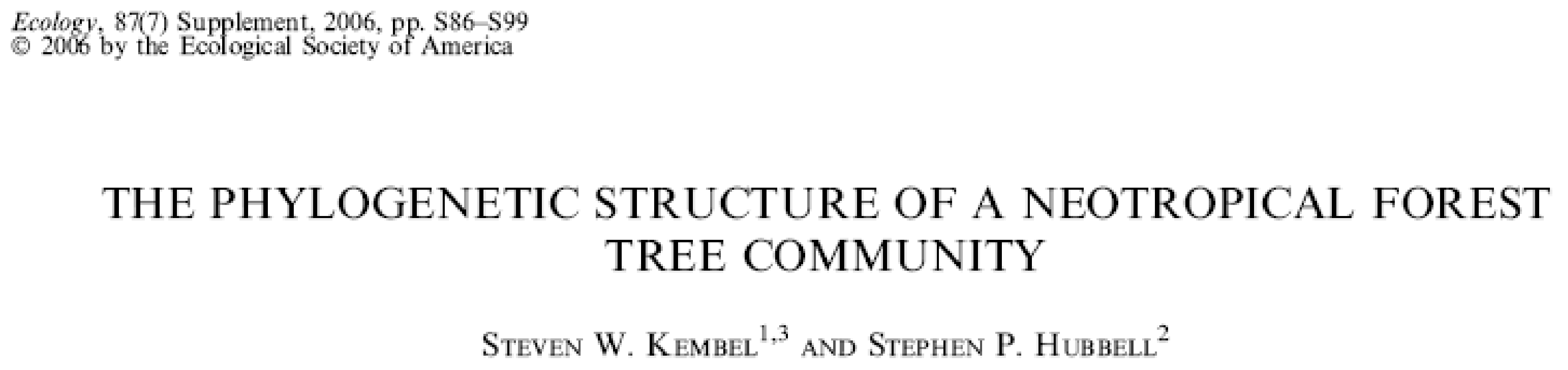


Figure 1 Schematic summary of various approaches to the integration of phylogeny, traits, and communities. (1) Examining the phylogenetic structure of community assemblages; (2) exploring the phylogenetic basis of community niche structure; (3) adding community context to studies of trait evolution and biogeography.

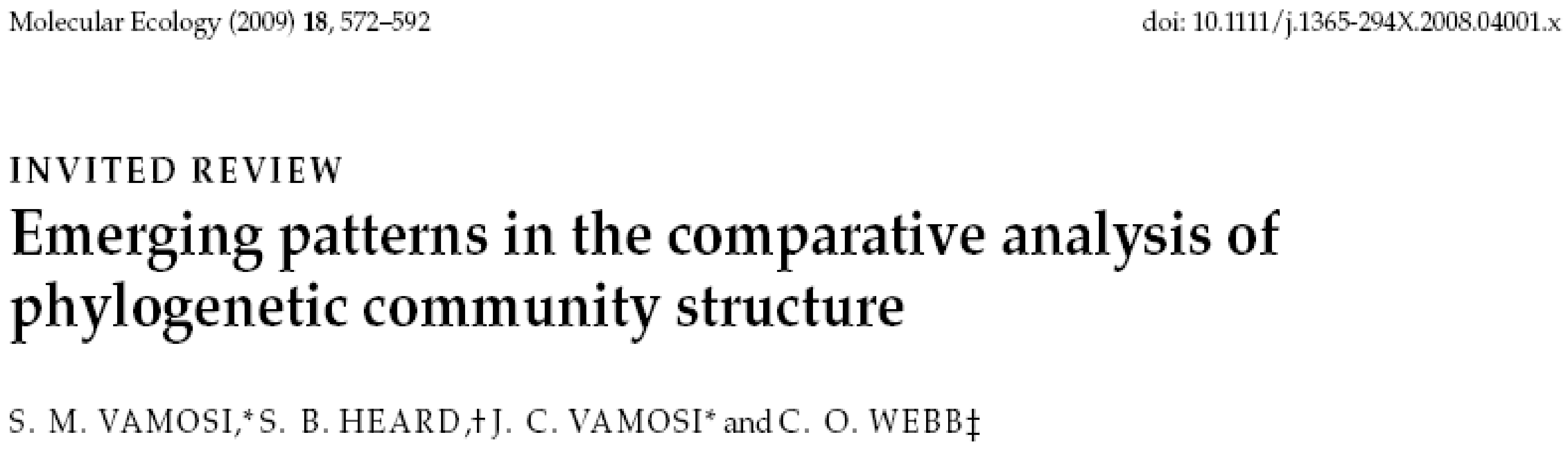
Webb et al. ARES 2002



Ecology, 87(7) Supplement, 2006, pp. S86–S99

THE PHYLOGENETIC STRUCTURE OF A NEOTROPICAL FOREST TREE COMMUNITY

Steven W. Kembel[1,3] and Stephen P. Hubbell[2]



Molecular Ecology (2009) **18**, 572–592 doi: 10.1111/j.1365-294X.2008.04001.x

INVITED REVIEW

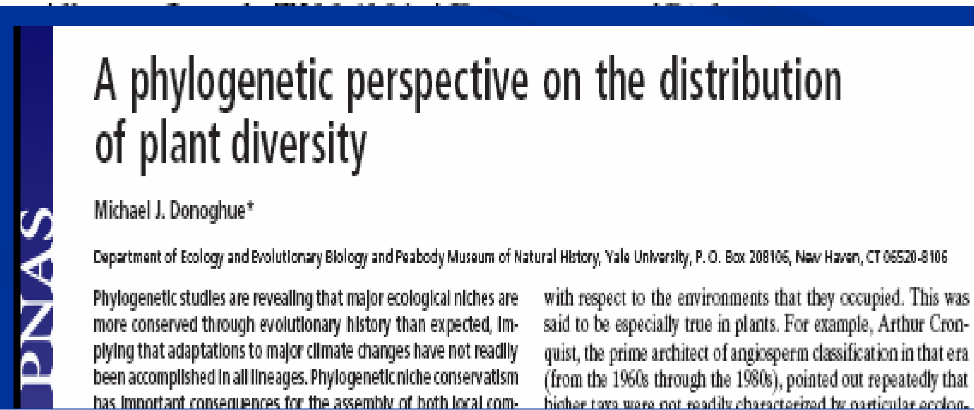
Emerging patterns in the comparative analysis of phylogenetic community structure

S. M. VAMOSI,* S. B. HEARD,† J. C. VAMOSI* and C. O. WEBB‡

Ecology Letters, (2009) **12**: 693–715 doi: 10.1111/j.1461-0248.2009.01314.x

REVIEW AND SYNTHESIS

The merging of community ecology and phylogenetic biology



PNAS

A phylogenetic perspective on the distribution of plant diversity

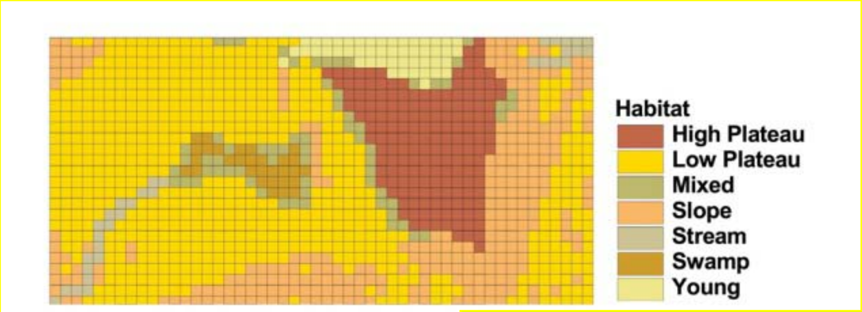
Michael J. Donoghue*

Department of Ecology and Evolutionary Biology and Peabody Museum of Natural History, Yale University, P. O. Box 208106, New Haven, CT 06520-8106

Phylogenetic studies are revealing that major ecological niches are more conserved through evolutionary history than expected, implying that adaptations to major climate changes have not readily been accomplished in all lineages. Phylogenetic niche conservatism has important consequences for the assembly of both local com-

with respect to the environments that they occupied. This was said to be especially true in plants. For example, Arthur Cronquist, the prime architect of angiosperm classification in that era (from the 1960s through the 1980s), pointed out repeatedly that higher taxa were not readily characterized by particular ecolog-

Barcodes
vs. Phylomatic



50-ha Forest Dynamics Plot on BCI, Panama (281 species): Community Phylogeny using a Supermatrix Approach with *rbcL / trnH-psbA / matK*

Phylomatic Phylogeny:

Phylogenetically clustered **= High Plateau, Low Plateau and Young Habitats**

Phylogenetically Over-dispersed **= Swamp and Slope Habitats**

Phylogenetically Random **= Stream and Mixed Habitats**

Barcode Phylogeny:

Phylogenetically clustered **= Low Plateau and Slope Habitats**

Phylogenetically Over-dispersed **= High Plateau, Mixed and Young Habitats =**

Phylogenetically Random **= Stream and Swamp Habitats**

Outline

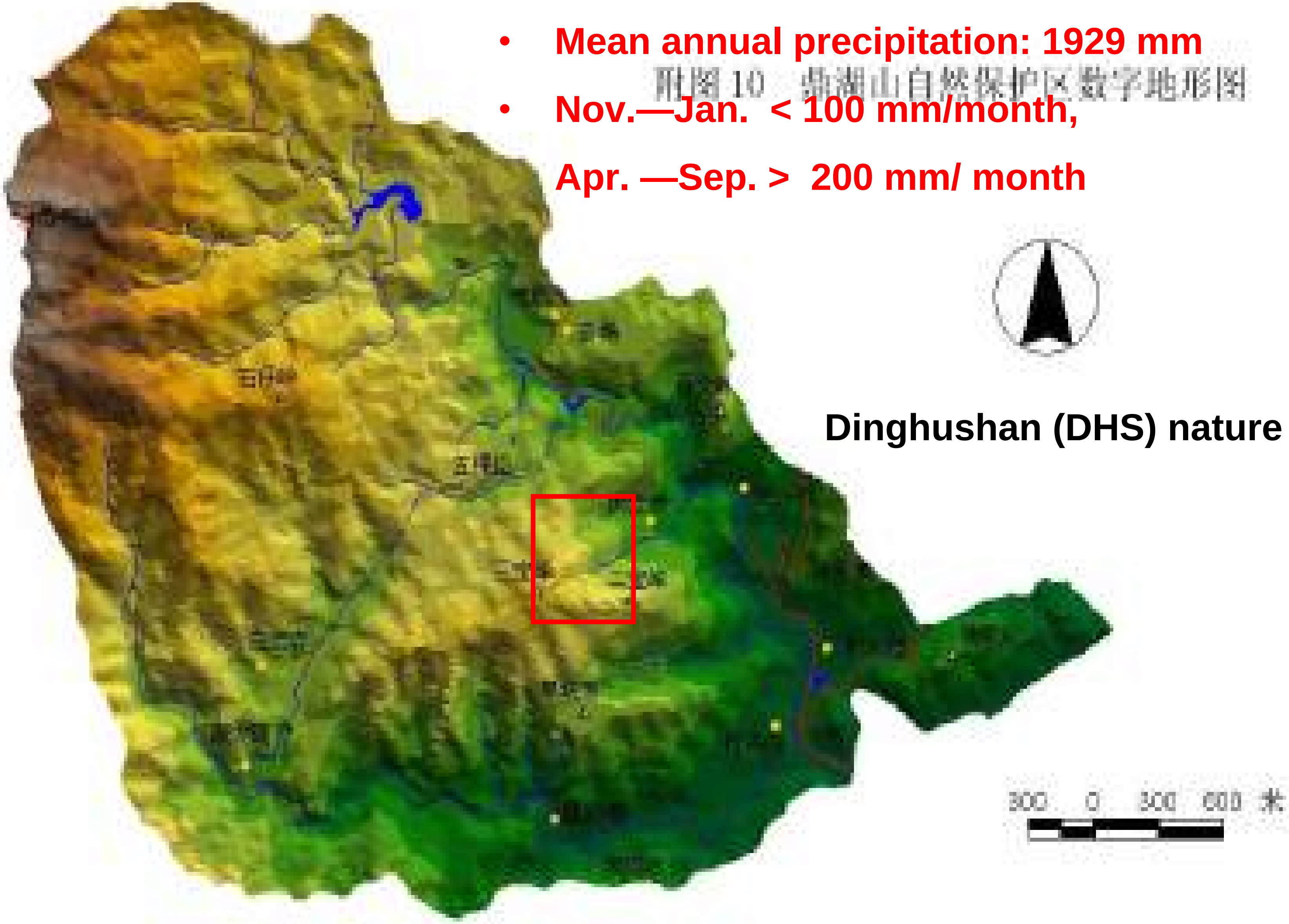
- Sampling
- Sequencing
- Tree reconstructing
- Community phylogenetic structure
- Conclusions

focus

The Center for Tropical Forest Science (CTFS)

- **Annual average temperature : 20.9 ℃**
- **Mean annual precipitation: 1929 mm**
- **Nov.—Jan. < 100 mm/month,**

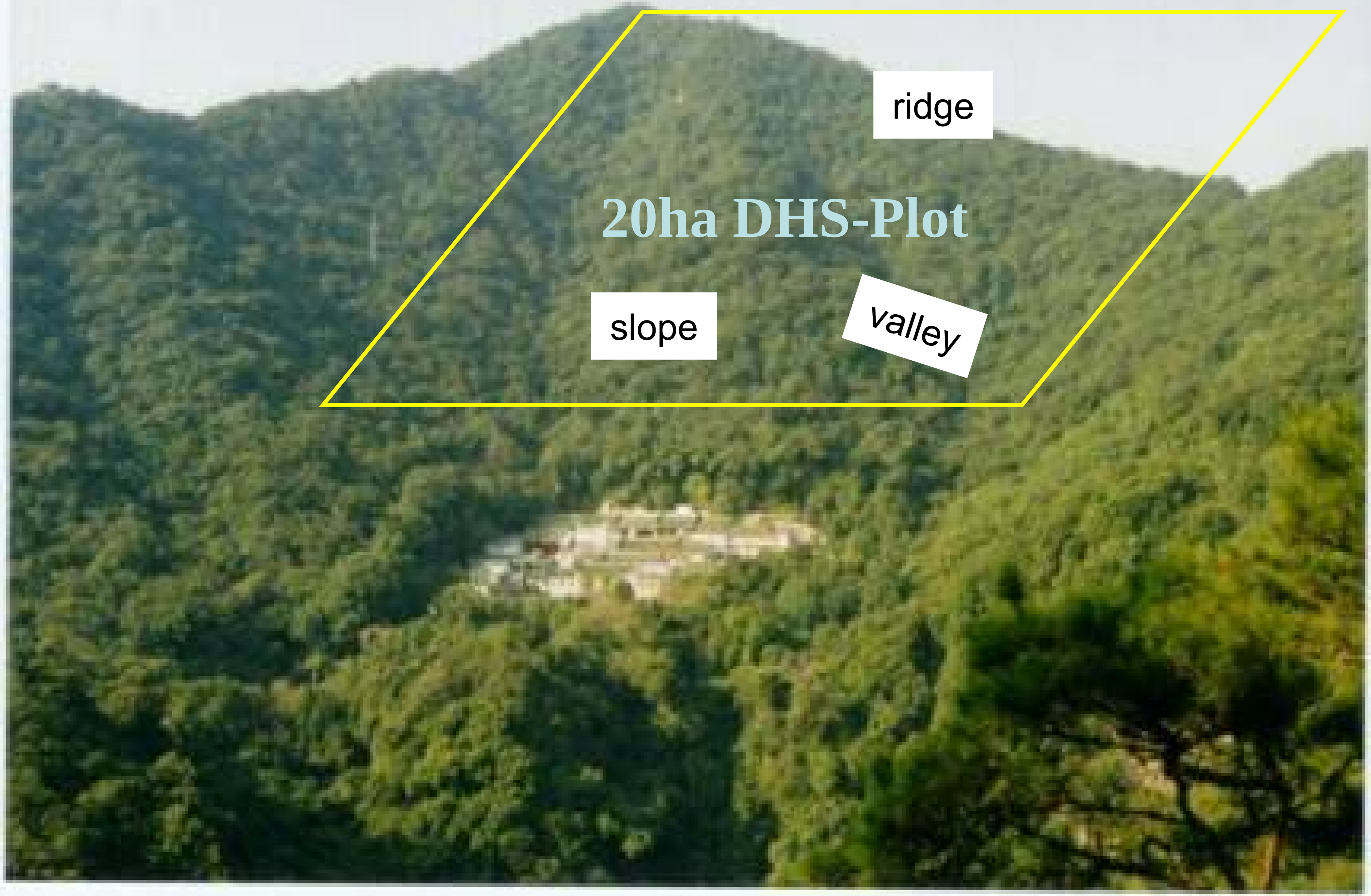
  Apr. —Sep. > 200 mm/ month



附图10 鼎湖山自然保护区数字地形图

Dinghushan (DHS) nature reserve

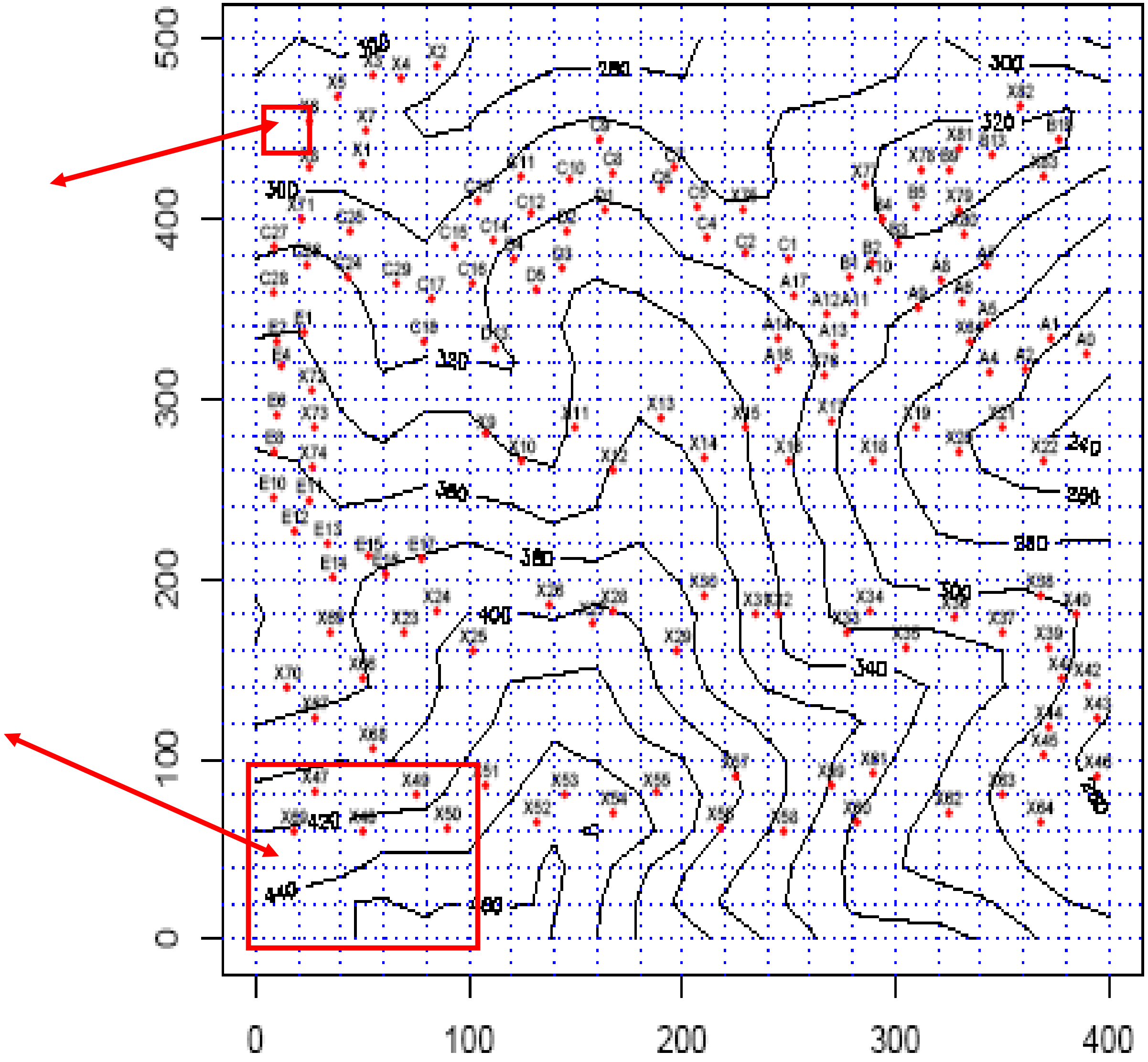
Different habitats



Different spatial scales

- 20*20 m,
500 subplots

- 100*100 m,
20 subplots



Sampling

- Species diversity in nature reserve:

278 families, more than 2000 species, including Gymnosperms, Angiosperms, Moss and Ferns

- Molecular sampling in the plot:

51 families

188 species

500 samples

Sequencing

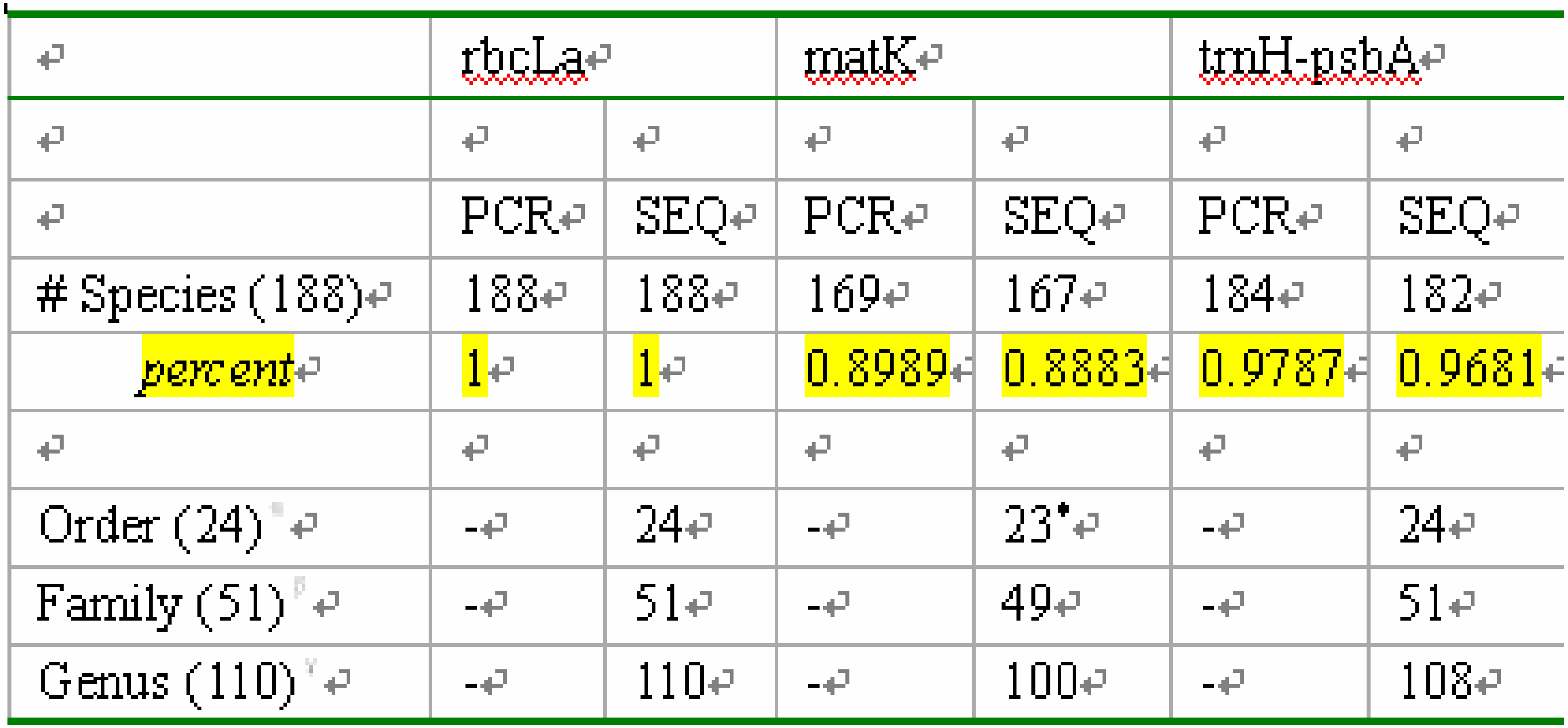
- A multi-locus barcode for land plants:

rbcLa

matK

trnH-psbA

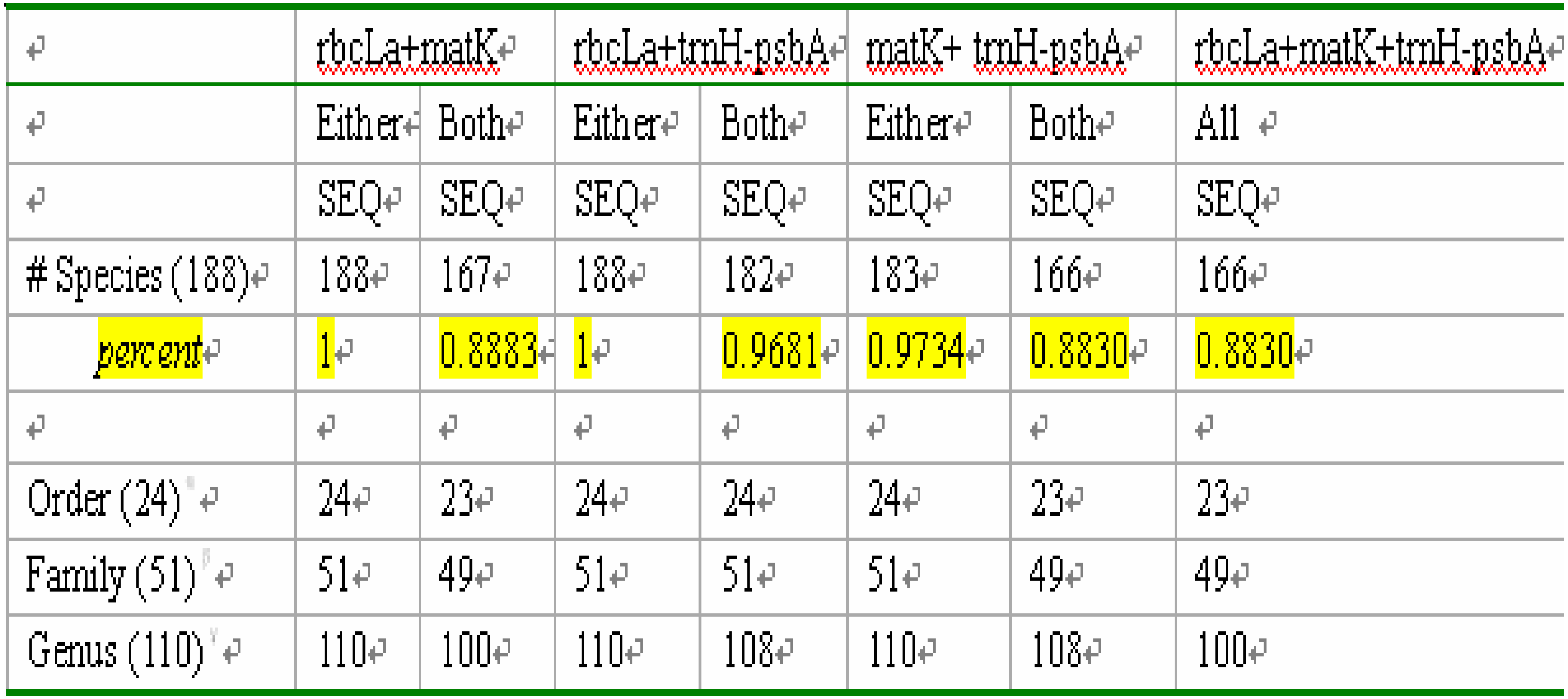
PCR and sequencing results of single primer



	rbcLa		matK		trnH-psbA	
	PCR	SEQ	PCR	SEQ	PCR	SEQ
# Species (188)	188	188	169	167	184	182
percent	1	1	0.8989	0.8883	0.9787	0.9681
Order (24)	-	24	-	23*	-	24
Family (51)	-	51	-	49	-	51
Genus (110)	-	110	-	100	-	108

rbcLa >trnH-psbA >matK

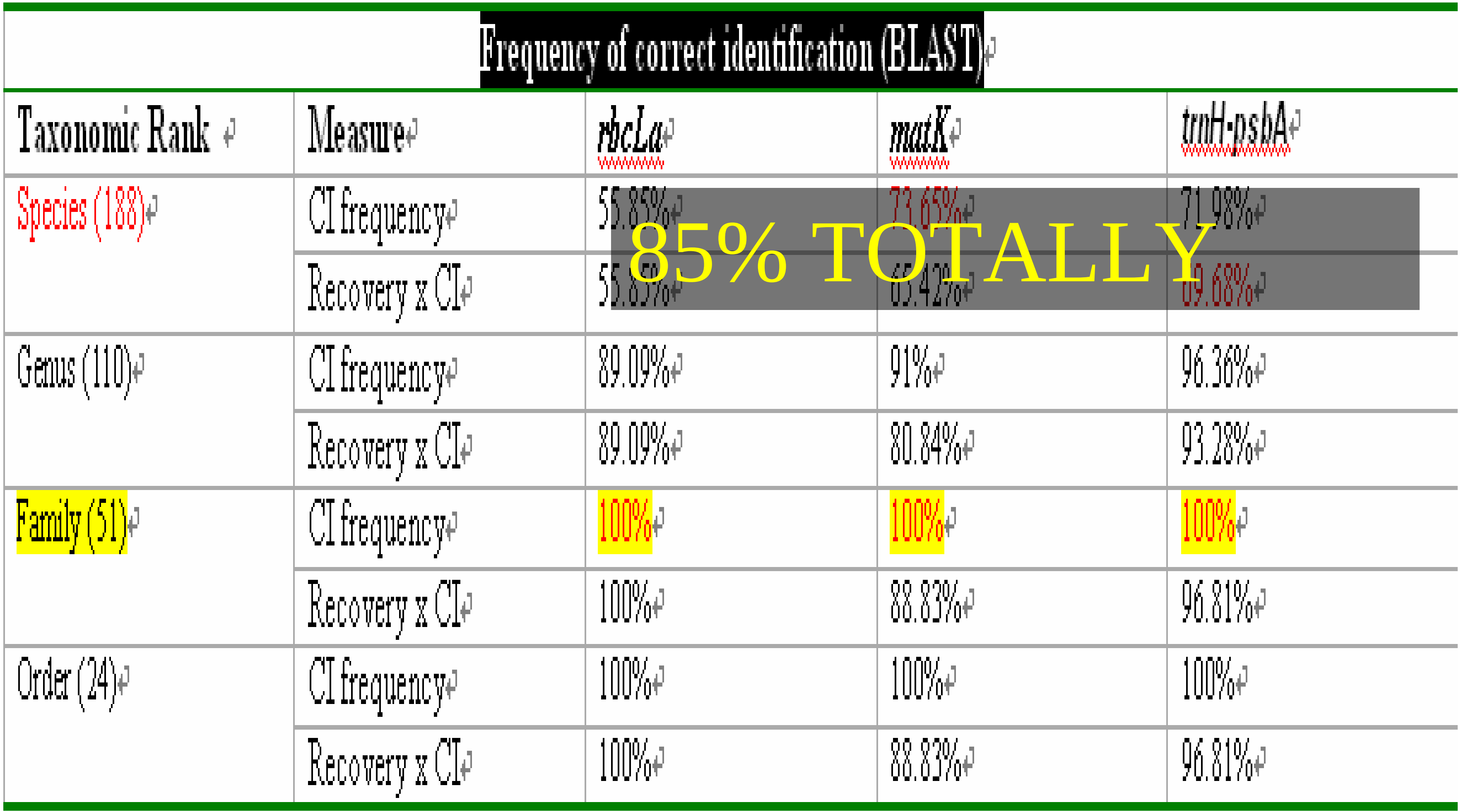
Sequencing results of combining primers



	rbcLa+matK		rbcLa+trnH-psbA		matK+ trnH-psbA		rbcLa+matK+trnH-psbA
	Either	Both	Either	Both	Either	Both	All
	SEQ	SEQ	SEQ	SEQ	SEQ	SEQ	SEQ
# Species (188)	188	167	188	182	183	166	166
percent	1	0.8883	1	0.9681	0.9734	0.8830	0.8830
Order (24)	24	23	24	24	24	23	23
Family (51)	51	49	51	51	51	49	49
Genus (110)	110	100	110	108	110	108	100

rbcLa+matK+trnH-psbA can provide the most information sites.

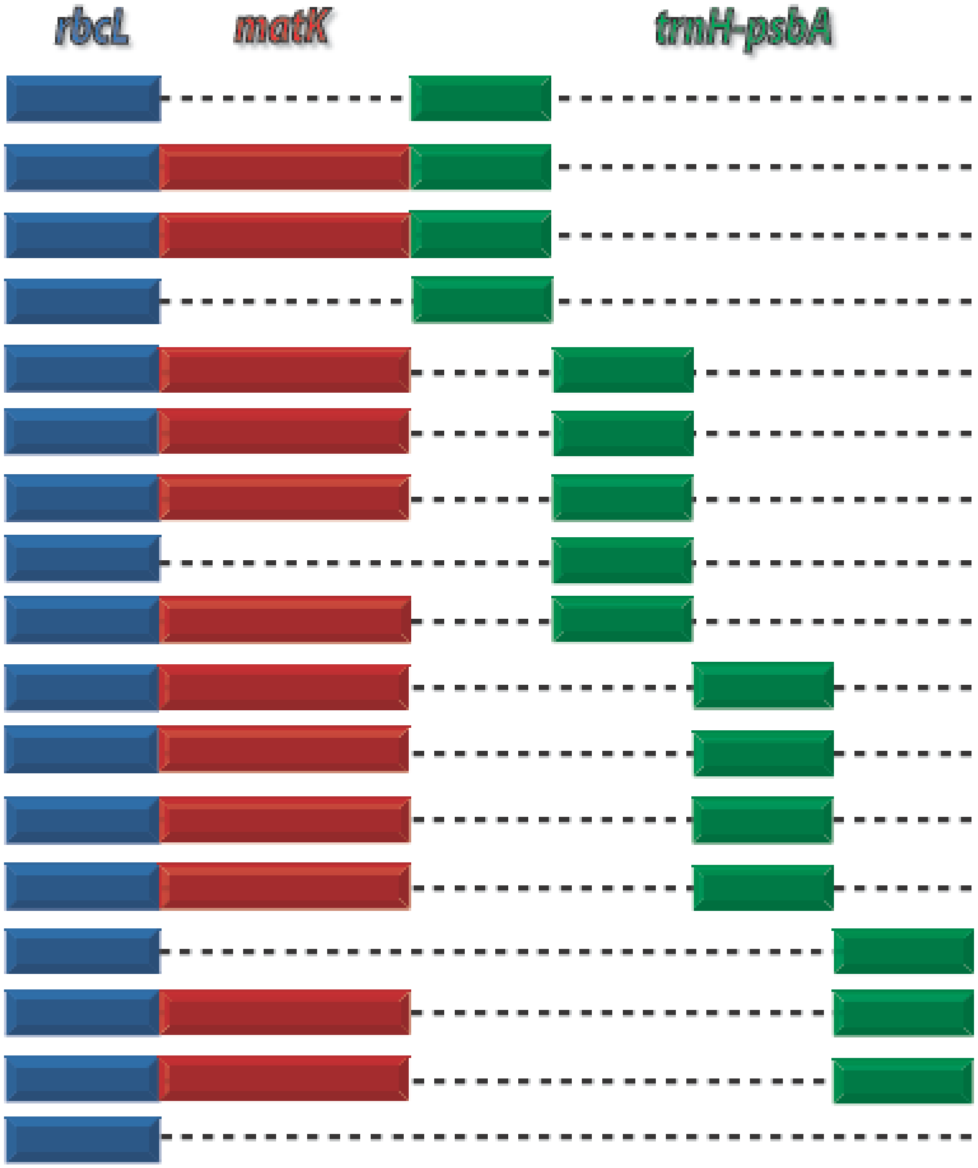
BLAST results for rates of correct identification (CI) for all species, as well as for genera, families and orders



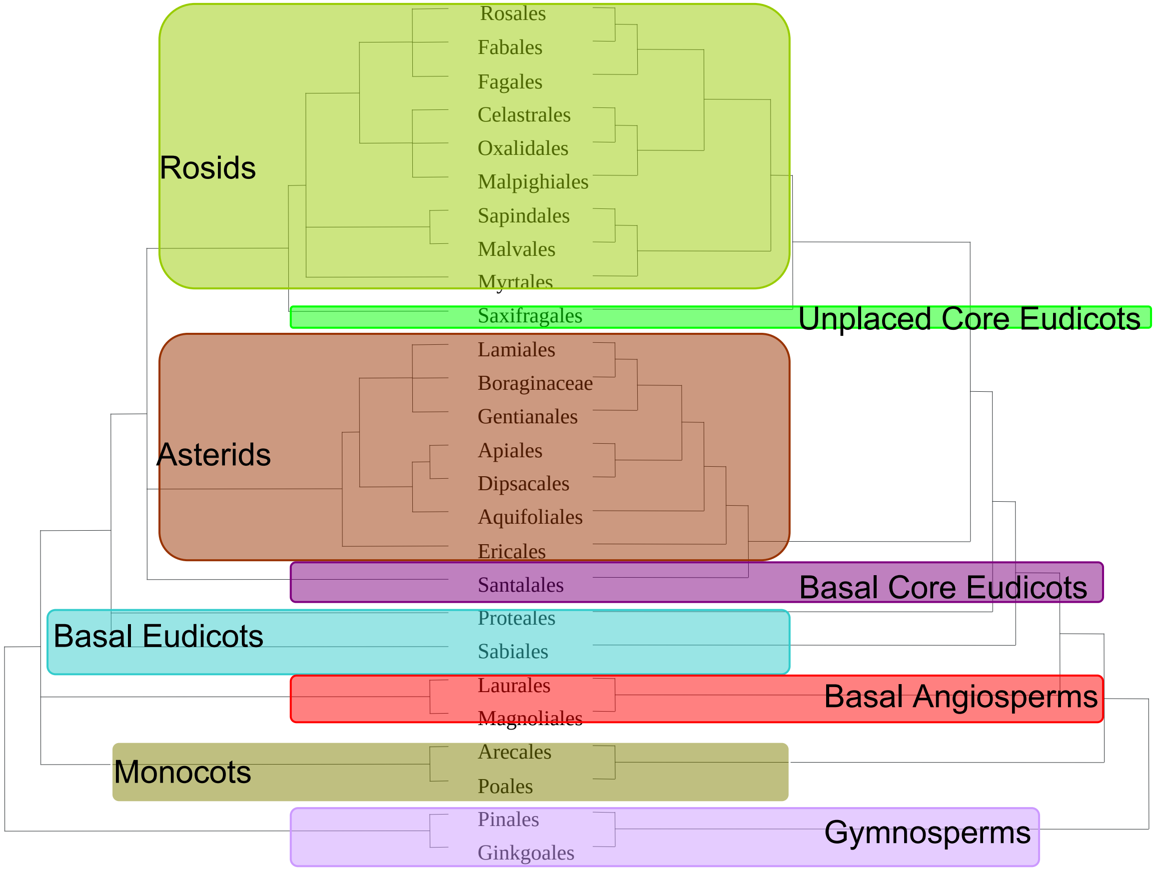
Frequency of correct identification (BLAST)				
Taxonomic Rank	Measure	*rbcLa*	*matK*	*trnH-psbA*
Species (188)	CI frequency	55.85%	73.65%	71.98%
	Recovery x CI	55.85%	65.42%	69.68%
Genus (110)	CI frequency	89.09%	91%	96.36%
	Recovery x CI	89.09%	80.84%	93.28%
Family (51)	CI frequency	100%	100%	100%
	Recovery x CI	100%	88.83%	96.81%
Order (24)	CI frequency	100%	100%	100%
	Recovery x CI	100%	88.83%	96.81%

85% TOTALLY

Barcode sequence supermatrix



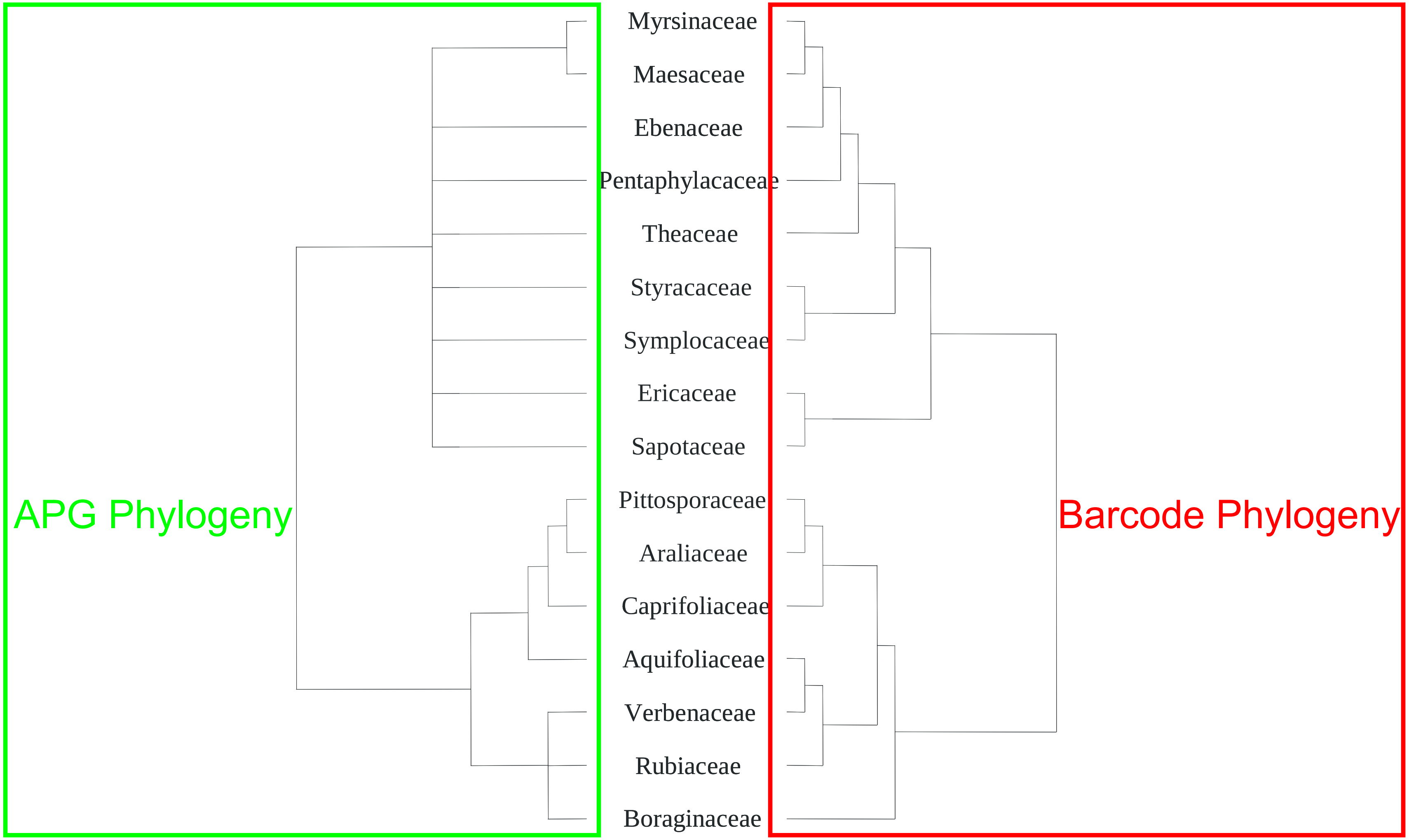
Barcode and APG Phylogeny of 26 orders in DHS plot



APG Phylogeny DHS Phylogeny

DHS phylogeny is better resolved than APG phylogeny.

Topology of Families within the Asterids of DHS plot



APG Phylogeny

Myrsinaceae
Maesaceae
Ebenaceae
Pentaphylacaceae
Theaceae
Styracaceae
Symplocaceae
Ericaceae
Sapotaceae
Pittosporaceae
Araliaceae
Caprifoliaceae
Aquifoliaceae
Verbenaceae
Rubiaceae
Boraginaceae

Barcode Phylogeny

Barcode phylogeny is better resolved than APG phylogeny.

Community assembly

Webb hypothesis:

- ►Small plot (local community), clustered?
- ►Large plot (pool community), overdispersed?

(Webb et al. 2000, 2002)

NRI & NTI

- Two index describing the community assembly

$$NRI = 1 - \left(\bar{x} \big/ \bar{x}_{\max}\right)$$

$$NTI = 1 - \left[\bar{x}_{\min} \big/ \left(\bar{x}_{\min}\right)_{\max}\right]$$

- Positive values >>> clustered
- Negative values >>> overdispersed

Local community vs. Pool community

- H_0: species are assembled into local communities at random

- For 20 m × 20 m

mean NRI=0.1411±0.0533, p=0.008

mean NTI=-0.0079±0.0470, p=0.866

Phylogenetic clustered in local community

- For 100 m × 100 m

mean NRI= 0.0711±0.2466, p=0.776

mean NTI=0.1568±0.2047, p=0.453

Random in pool community

Habitat partitioning hypothesis

- Habitat heterogeneity

important in the maintenance of tropical forest tree species diversity

Five habitats in DHS plot

- high/low slope
- high/low valley
- ridge

Community assembly patterns

Phylomatic Phylogeny:

Phylogenetically clustered =Low Slope and Low Valley Habitats

Phylogenetically Over-dispersed = High Slope, High Valley and Ridge Habitats

Barcode Phylogeny:

Phylogenetically clustered = Low Slpoe and Low Valley Habitats

Phylogenetically Over-dispersed = High Slope and Ridge Habitats

Phylogenetically Random = High Valley Habitat

Some differentiations

- Correct barcoding identification:

98% in BCI using three-locus barcodes;

96% in French Guiana using morphology and molecular, but <70% using any barcode

- Community phylogenetic structure (BCI):

phylogenetic clustering: high plateau, young and young habitats;

phylogenetic overdispersion: swamp and slope habitats

Conclusions

- Three-locus barcode well identify all taxa
- Plant DNA barcodes well resolve the phylogeny of Dinghushan forest community
- *rbcLa+matK+trnH-psbA* supermatrix performs well in tree reconstructing
- Species in local community were more phylogenetically related than expected by chance, while in pool community did not
- Different habitat shows different community phylogenetic structure
- Highly resolved phylogenies based on DNA barcode sequence data will enhance research focused on the interface between community ecology and evolution

☺Thank you for attention!

Collaborators:
Professor Wan-Hui Ye
Doctor Ju-Yu Lian
Professor John Kress
Doctor David Erickson
et al.