

**Phylogenomics  
of the  
Brassicales  
Brassicaceae  
and  
Tribe  
Brassiceae**

**J. Chris Pires**

**University of Missouri**

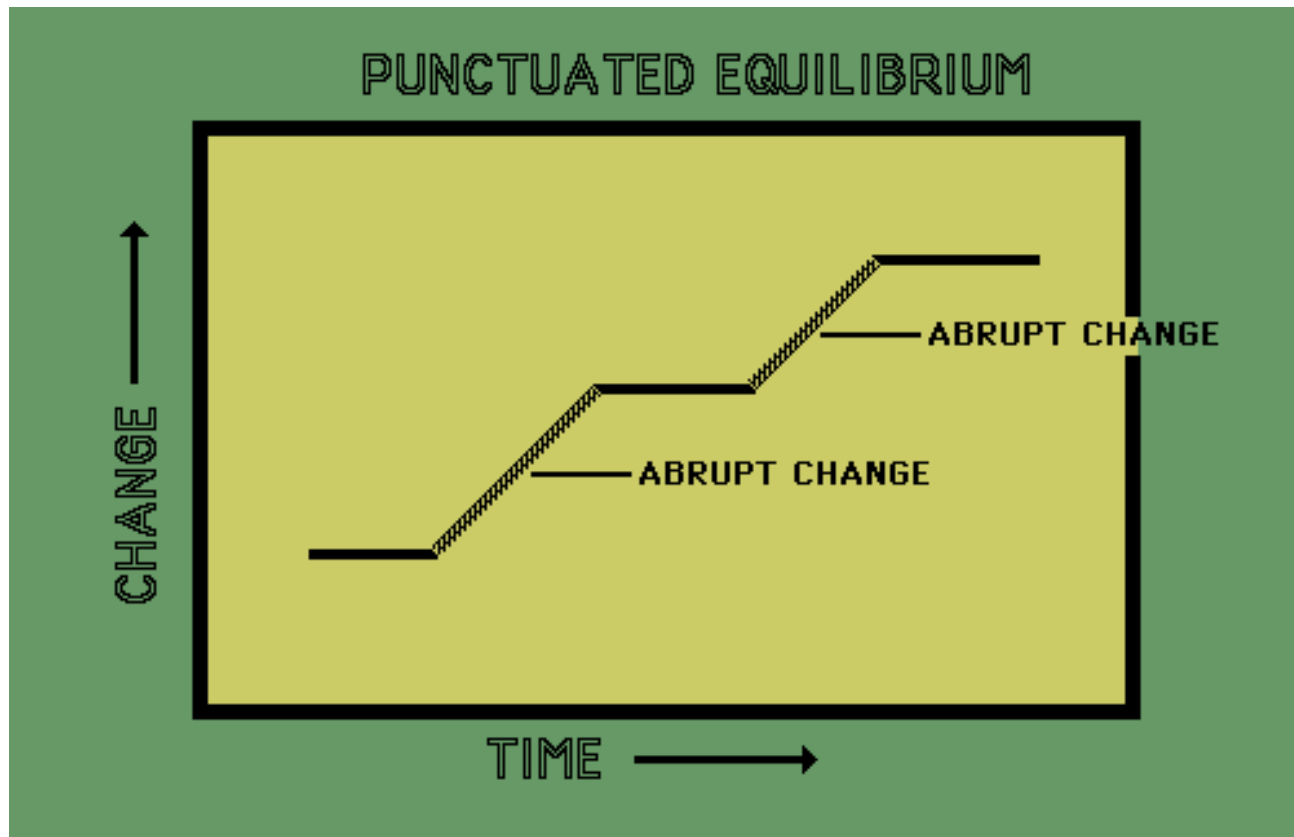
**31 March 2014**



**@JChrisPires**



# Polyploid evolution and punctuated equilibrium: dynamic periods of change followed by stasis?



Are whole genome duplications (WGD) associated with morphological complexity and/or clade diversification?

# Outline: Phylogenomics

**I. Order Brassicales**

**II. Family Brassicaceae**

**III. Tribe Brassiceae**

**IV. Future Directions**

# Order Brassicales Questions

Does polyploidy spur lineage diversification and complexity?  
Specifically, have nested polyploid events contributed to the diversification of lineages and anti-insectivore secondary metabolites (glucosinolates)? Revisit Ehrlich and Raven co-evolution/macroevolution ideas:

A. PHYLOGENOMICS – need trees to interpret WGD

B. TRANSCRIPTOMICS – need to find WGD



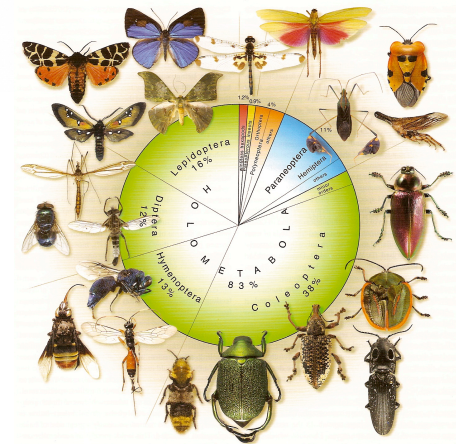
C. NETWORK BIOLOGY – measure WGD functional impact  
On origin of novel traits and diversification

# Phylogenomics using Illumina Sequencing

## Genome Survey Sequencing (GSS)

1. Total DNA or mRNA
  2. Multiplex Libraries
  3. SE or PE reads
  4. de novo Assembly (cp, mt, rDNA)
  5. Annotation
  6. Phylogenetics Analyses
- (\$100/taxa; **Steele et al. 2012 AJB**)

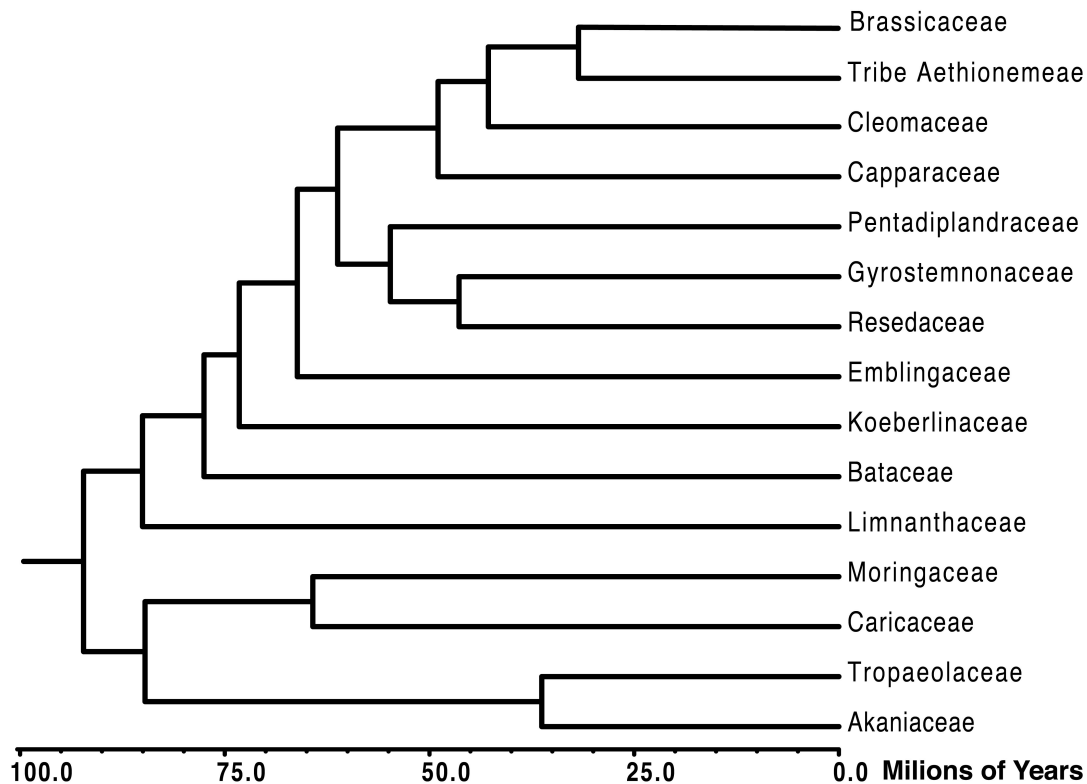
Time Travel !



1.6. The diversity of recent hexapods as proportions of named species.

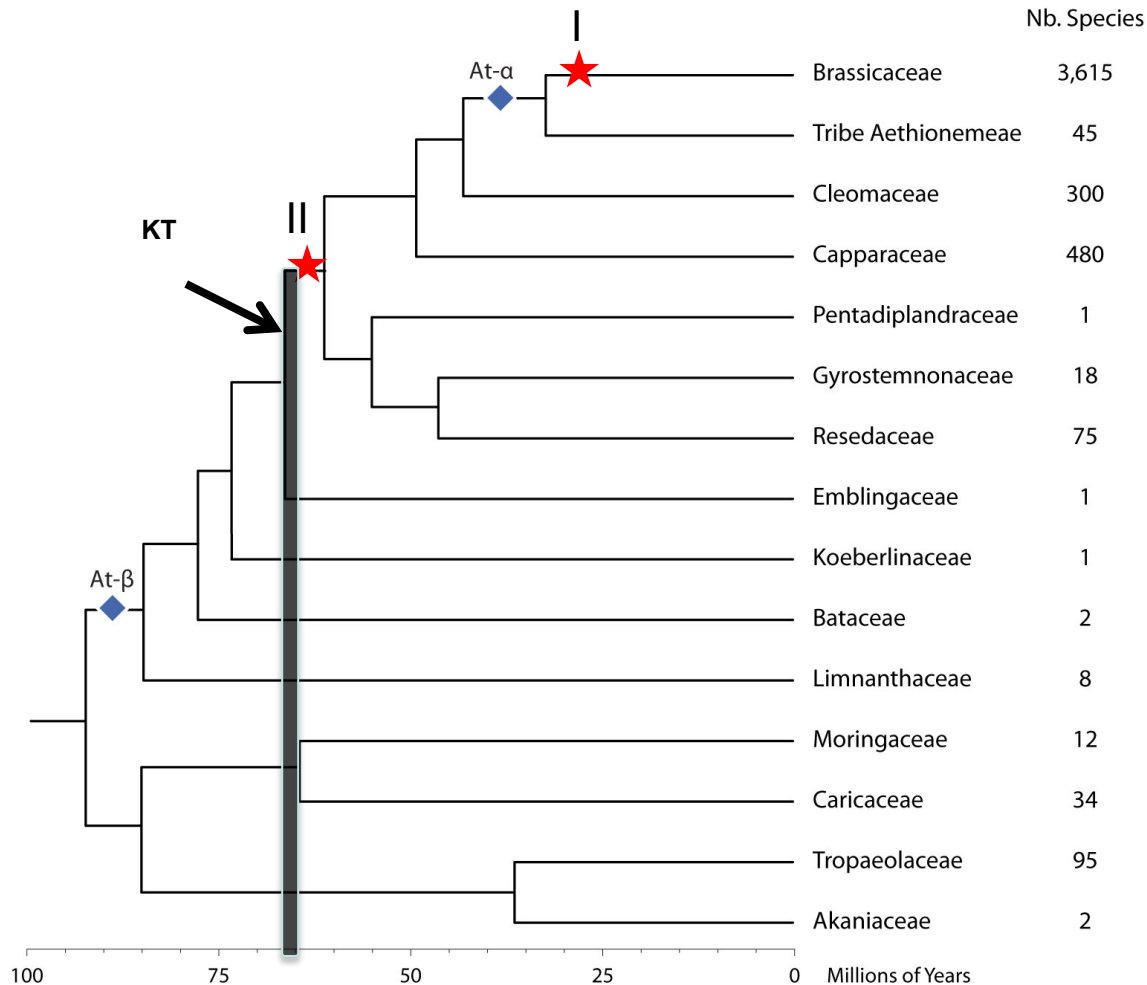
Species	Illumina Read Processing		Nuclear Genome		Mitochondrial Genome		Plastid Genome	
	Raw Reads	Cleaned Reads	Total Contigs	Total Sequence	Total Contigs	Total Sequence	Total Contigs	Total Sequence
<i>Myagrum</i>	14,299,919 Reads	13,621,738 Reads	22,569 Contigs 8,139bp - 100bp	4,003,494 bp	26 Contigs 45,633bp - 108bp	271,921 bp	6 Contigs 46,452bp - 5,524bp	127,770 bp
<i>Caulanthus</i>	13,756,796 Reads	13,035,325 Reads	8,636 Contigs 10,337bp - 100bp	2,452,489 bp	31 Contigs 39,440bp - 120bp	296,933 bp	3 Contigs 92,762bp - 12,993bp	126,303 bp
<i>Ionopsidium</i>	14,201,297 Reads	13,436,285 Reads	9,487 Contigs 10,928bp - 100bp	2,569,346 bp	125 Contigs 37,317bp - 104bp	515,028 bp	6 Contigs 50,268bp - 5,441bp	149,512 bp
<i>Heliophila</i>	12,084,646 Reads	11,434,650 Reads	5,688 Contigs 8,201bp - 100bp	1,470,345 bp	27 Contigs 36,875bp - 127bp	319,313 bp	4 Contigs 46,202bp - 5,683bp	125,890 bp
<b>Average</b>	13,585,665 Reads	12,882,000 Reads	11,595 Contigs 9,401bp - 100bp	2,623,919 bp	52 Contigs 39,816bp - 115bp	350,799 bp	5 Contigs 58,921bp - 7,410bp	132,369 bp

# Brassicales: 18 families (4765 species)



- **Phylogeny estimated with 1155 single copy nuclear genes**
- **Largely congruent with 78 plastid & 34 mitochondrial markers**
- **ML (RaxML), MP (PAUPRat), and Bayesian (BEAST)**
- **100% Bootstrap support**

# Diversification Analyses show a “Lag-Time” for Speciation after WGD



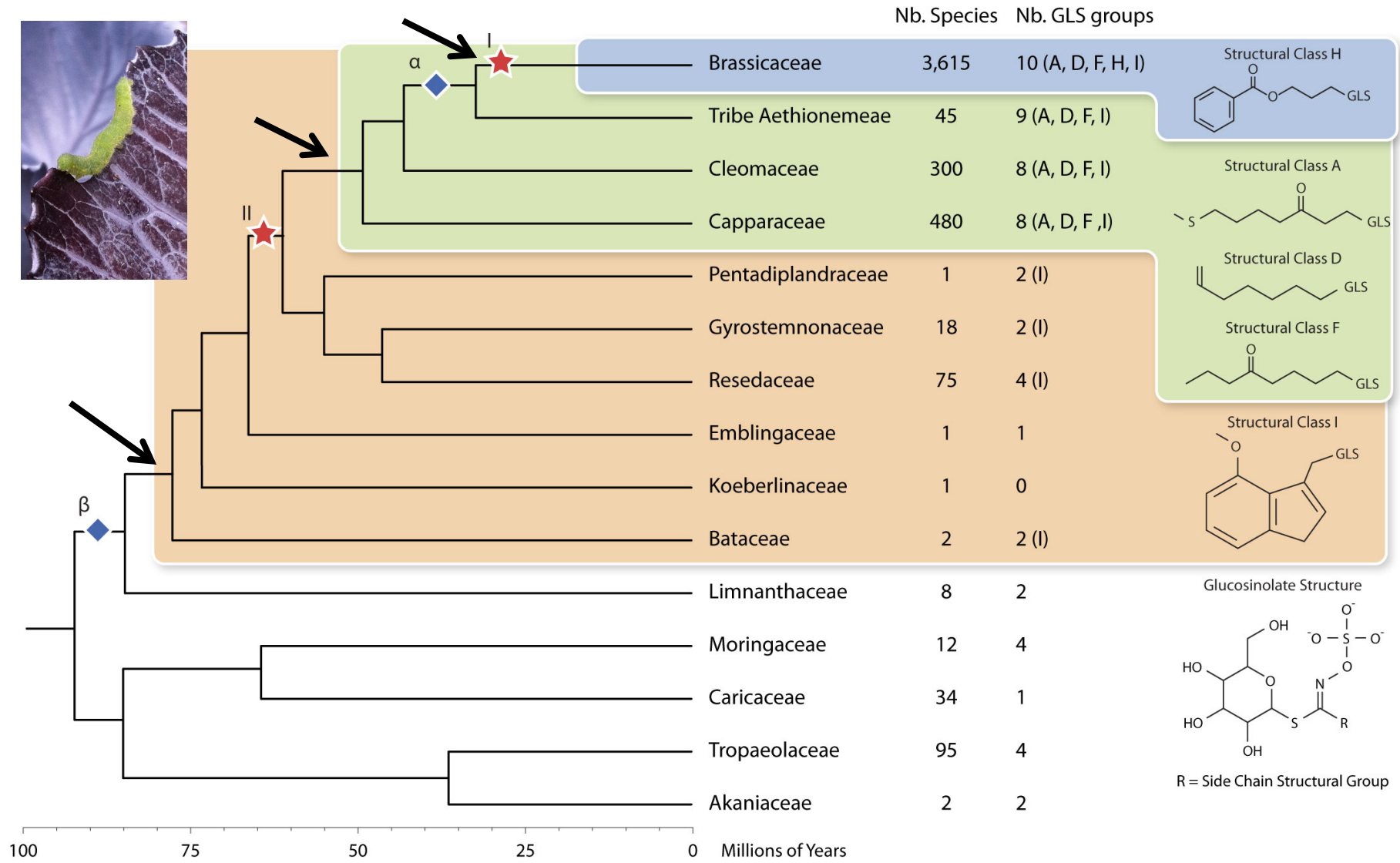
**I. Brassicaceae radiation:**  
 p-value = 0.0003430522  
 bonf. p-value = 0.006861044  
 “Mass-diversification”  
 following **At-α WGD**  
 occurred ~ 10 mya later

**II. KT-Extinction radiation:**  
 p-value = 0.0019349845  
 bonf. p-value = 0.038699690  
 “Mass-diversification”  
 following **At-β WGD**  
 occurred minimum  
 ~17 million years later

Thus, WGD alone does not lead to diversification

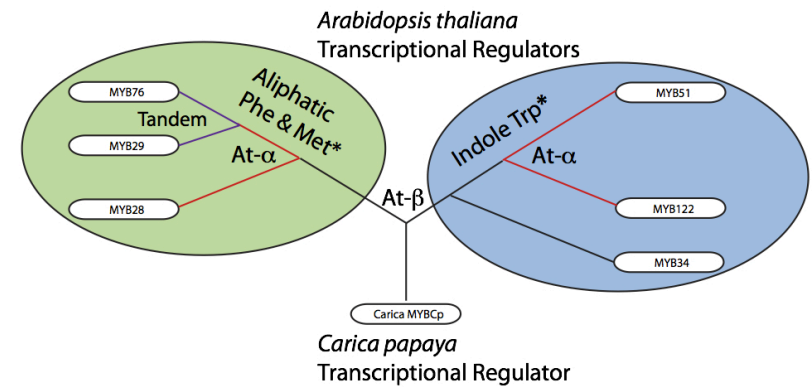
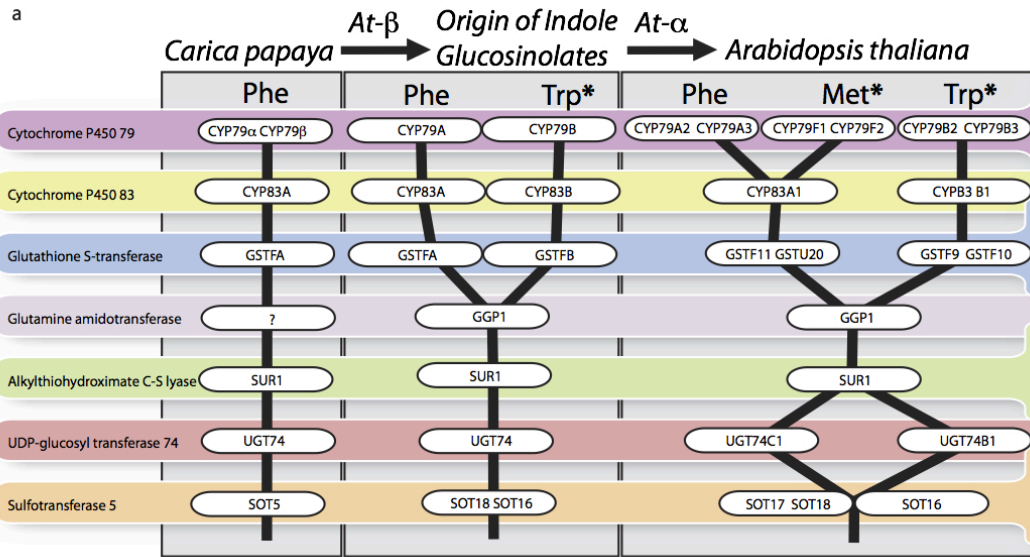
If lag, how many **species radiations** be associated with **WGD** ?

# Origin of Novel Chemical Defenses

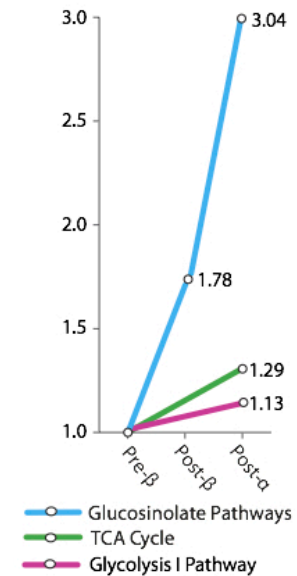
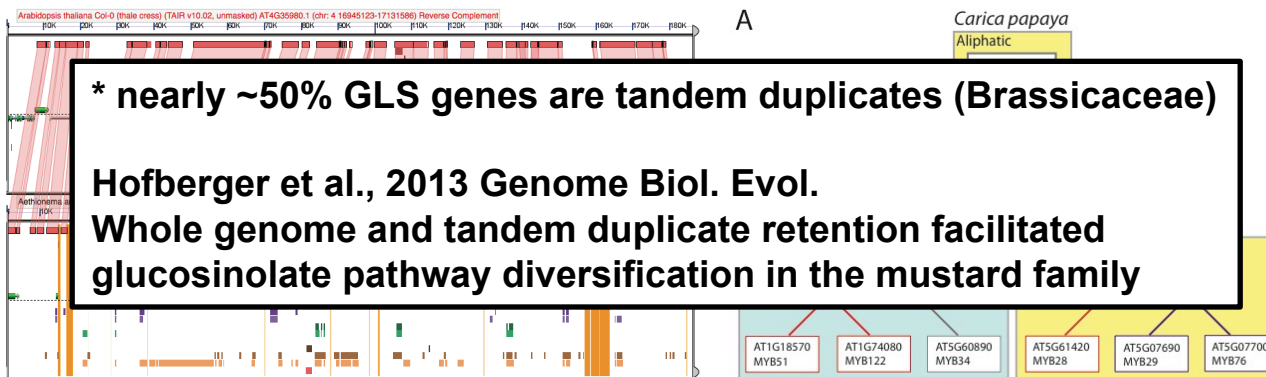




# Origin of Novel Pathways: WGDs



- Retained Whole Genome Duplicates (Neofunctionalized) At-β and At-α event – Novel Biosynthetic Pathways
- \* Supported by Syntenic Blocks and Phylogenetic Analyses



# Brassicales Summary

1. **The Butterfly Plant Evolutionary Arms Race has a new fossil-calibrated phylogenomic chronogram**
2. **Origin of Novel Glucosinolate Pathways arose due to WGD, the At- $\beta$  WGD spurred arms race with indole GLS, and likely At- $\alpha$  WGD rearmed for some met-GLS**
3. **Specialist insect herbivores turn the ‘mustard oil bomb’ into a defensive ‘cyanide bomb’**
4. **Significant Adaptive Radiations on both sides**
  - A. **Brassicales-feeding Pierid butterflies are 4X more species-rich**
  - B. **Brassicales plants with Indolic GLS are 29X more species-rich**
  - C. **Brassicaceae plants with Met GLS are 80X more species-rich**

# Outline: Phylogenomics

**I. Order Brassicales**

**II. Family Brassicaceae**

**III. Tribe Brassiceae**

**IV. Future Directions**

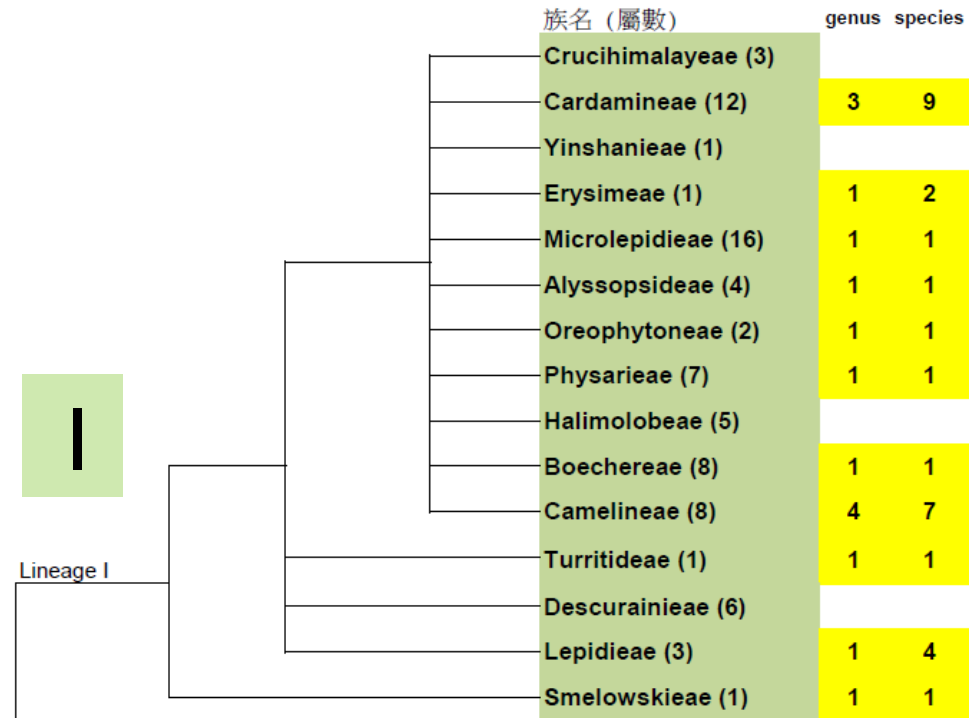
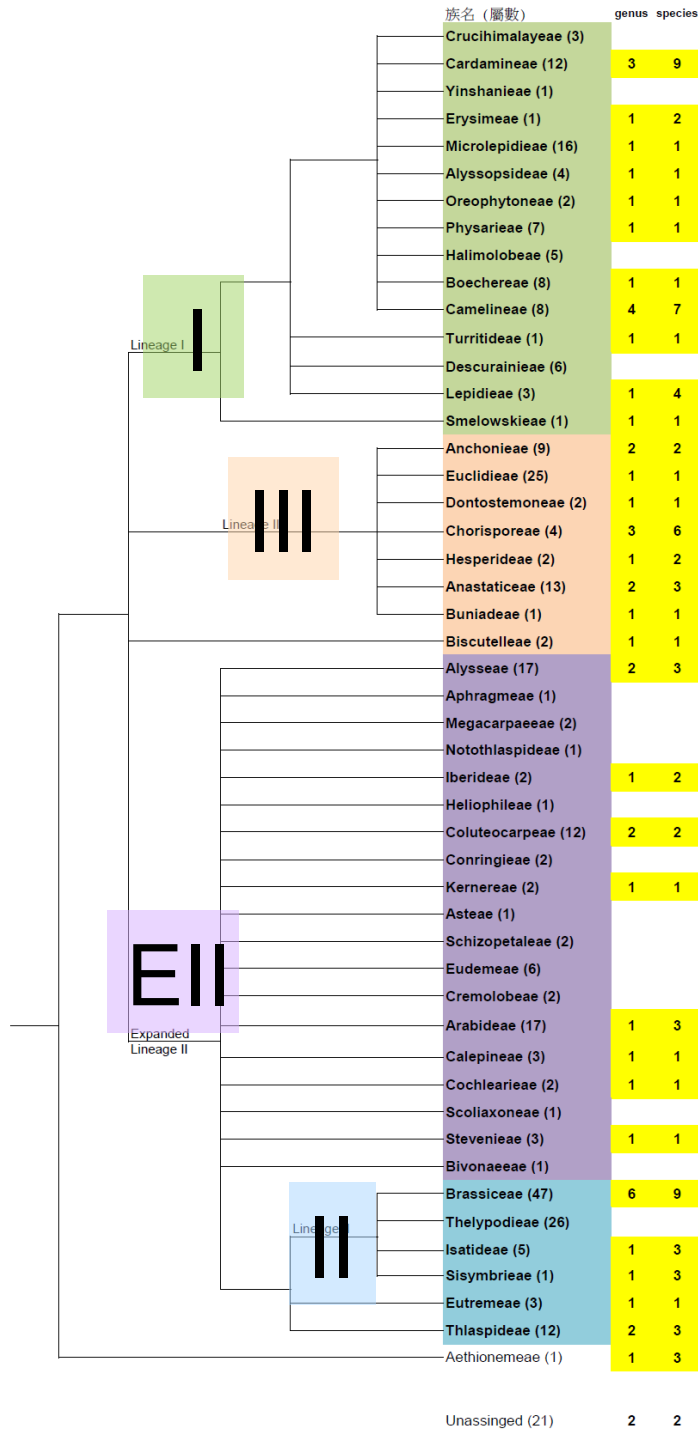
**Highly conserved low-copy nuclear genes as effective markers for phylogenetic analyses in family Brassicaceae**

**Collaborators: Chien-hsun Huang , Hong Ma  
Ihsan Al-Shebaz, Marcus Koch, Klaus  
Mummenhoff and others**

**Current taxon sampling: 85 species  
(across 33 tribes, 52 genera, 84 species)**

**Sequence data: 28 single copy nuclear genes  
from 6 genomes, 58 transcriptomes, and 27 PCR**

# Lineages I, II, III and tribes within the family Brassicaceae remain unresolved; taxon sampling shown in yellow



# Outline: Phylogenomics

**I. Order Brassicales**

**II. Family Brassicaceae**

**III. Tribe Brassiceae**

**IV. Future Directions**

# Plant materials



**Seed collection completed**

USDA, Madrid, Field trip



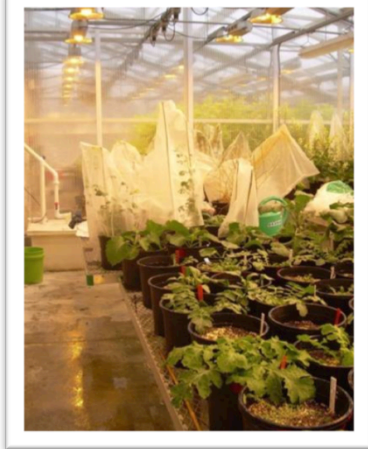
**Plants were grown greenhouse**

Life Sciences Center MU



**DNA and Herbarium vouchers**

Collections identified: Ihsan Al-Shebaz

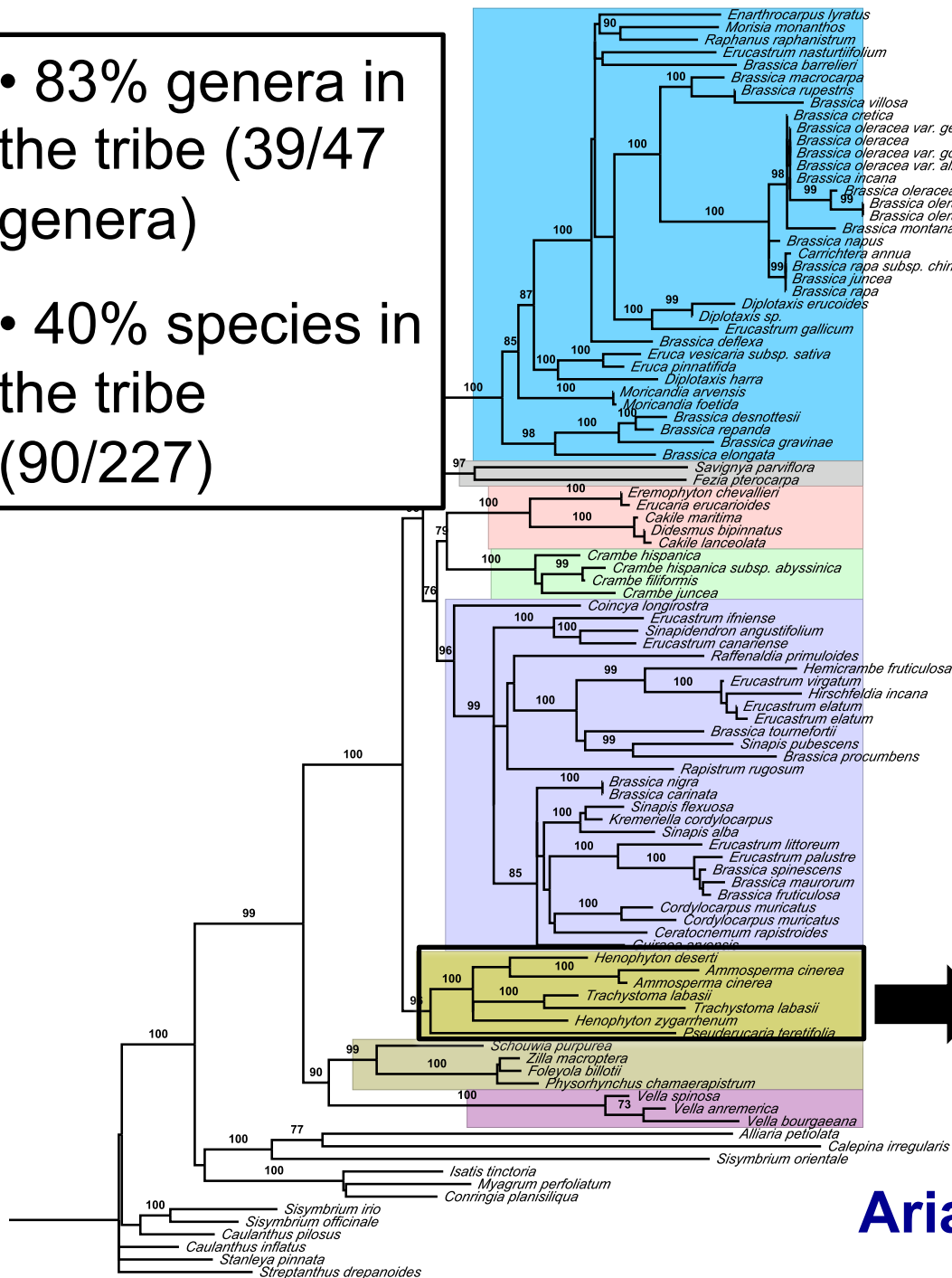


- 83% genera in the tribe (39/47 genera)
- 40% species in the tribe (90/227)

# Results

4-chloroplast marker phylogeny

1. Relationships among subtribe lineages in **tribe Brassiceae** more resolved
2. Recovered 7 previously identified clades and **New African clade**
3. Parallel evolution



**Arias & Pires 2012 Taxon**



**Tribe Brassiceae** has seven major lineages (or sub-tribes) identified using cp markers (Warwick 2005), but not well resolved

Oleracea Nigra Cakile Vella Zilla Savignya Crambe

**Genera**

10

15

4

4

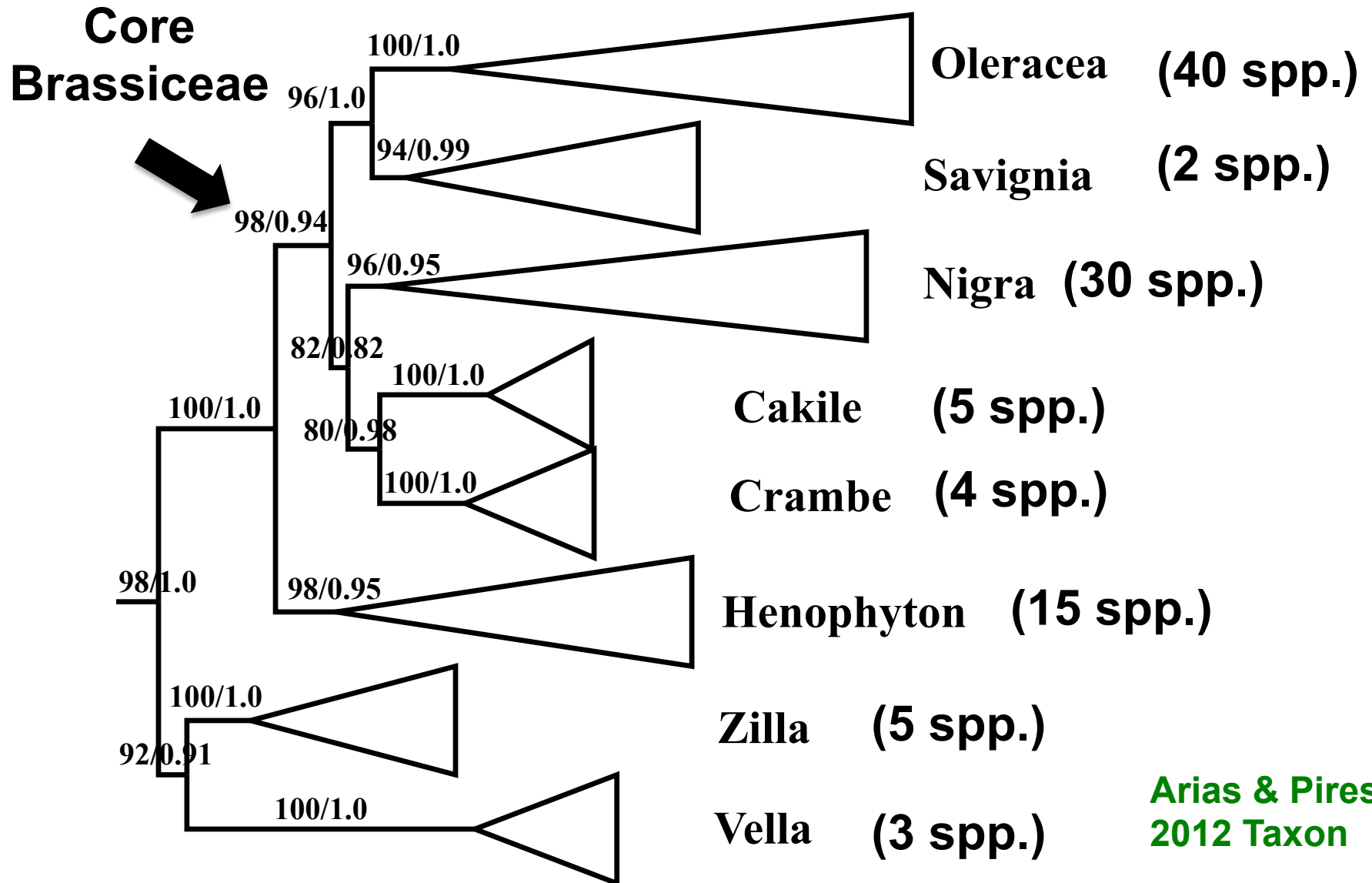
5

3

1



# Relationships among sub-tribes in Tribe Brassiceae are more resolved

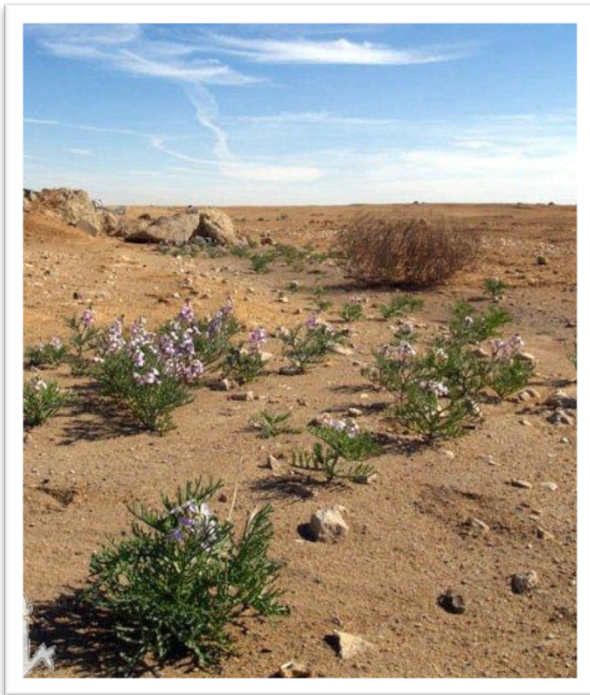


# Henophyton Lineage: New African clade

*Henophyton*  
*Trachystoma*  
*Ammosperma*  
*Pseuderucaria*

- Rare and Endemic genera from North Africa (Morocco and Algeria)

- DNA extracted from old herbarium specimens at MA



*Pseuderucaria*

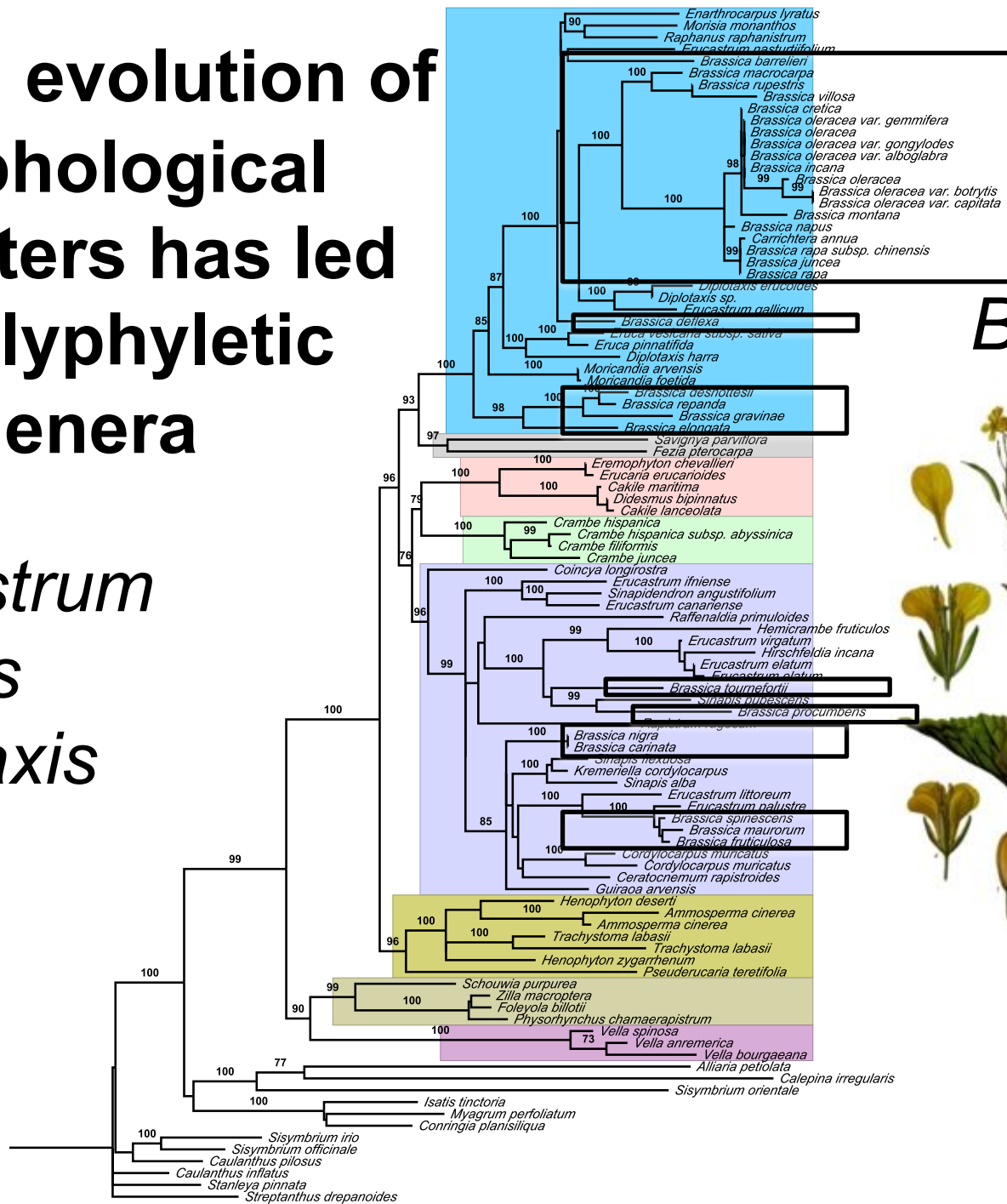


*Henophyton*

**Arias & Pires  
2012 Taxon**

# Parallel evolution of morphological characters has led to polyphyletic genera

*Erucastrum*  
*Sinapis*  
*Diplotaxis*



*Brassica*



## When did the tribe diversification occur?

The tribe origin & triplication event have been dated at

≈ 28 to 15.6 Mya (Beilstein & al., 2010)

≈ 23 to 11 Mya (Couvreur & al., 2010)

### Diversification times

R8s (Sanderson 2004) and BEAST (Drummond & Rambaut 2007)

### Diversification rates

symmeTREE (Moore et al. 2004)

### Biogeography

RASP (Yu et al. 2012 in review)

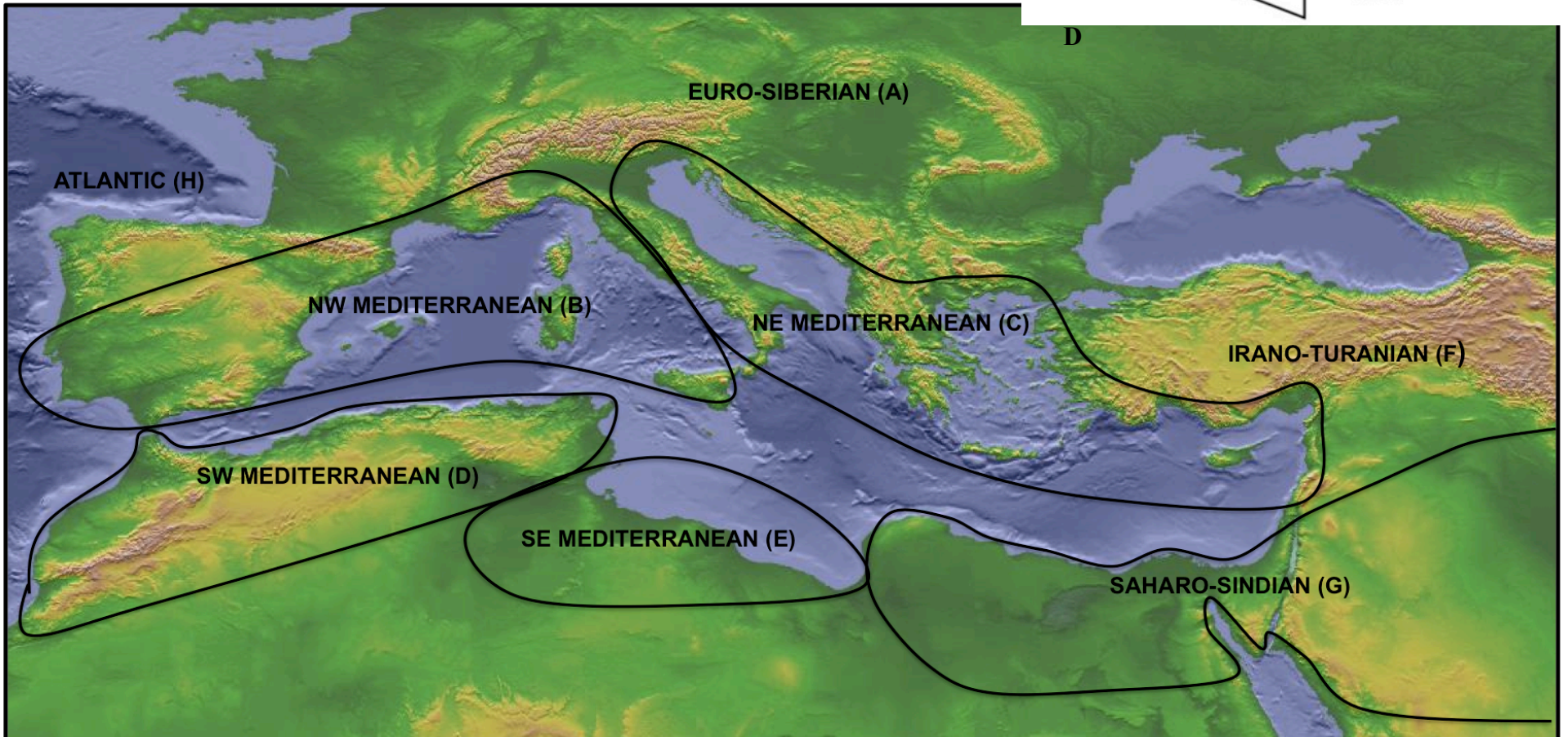
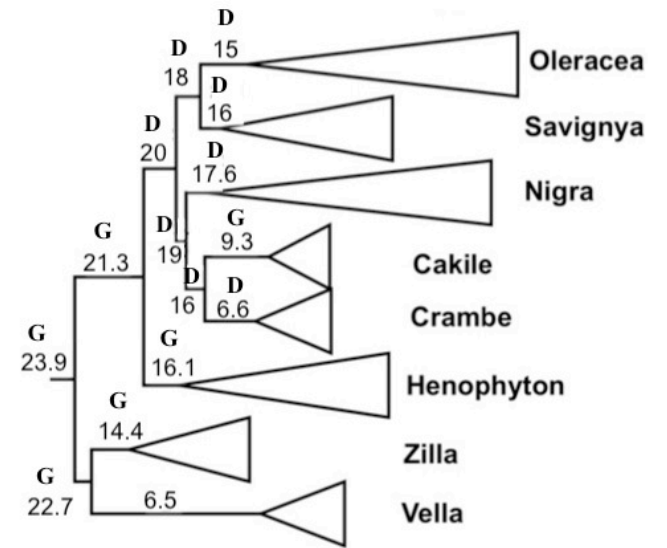
### Character state reconstructions

Mesquite (Maddison & Maddison 2006)

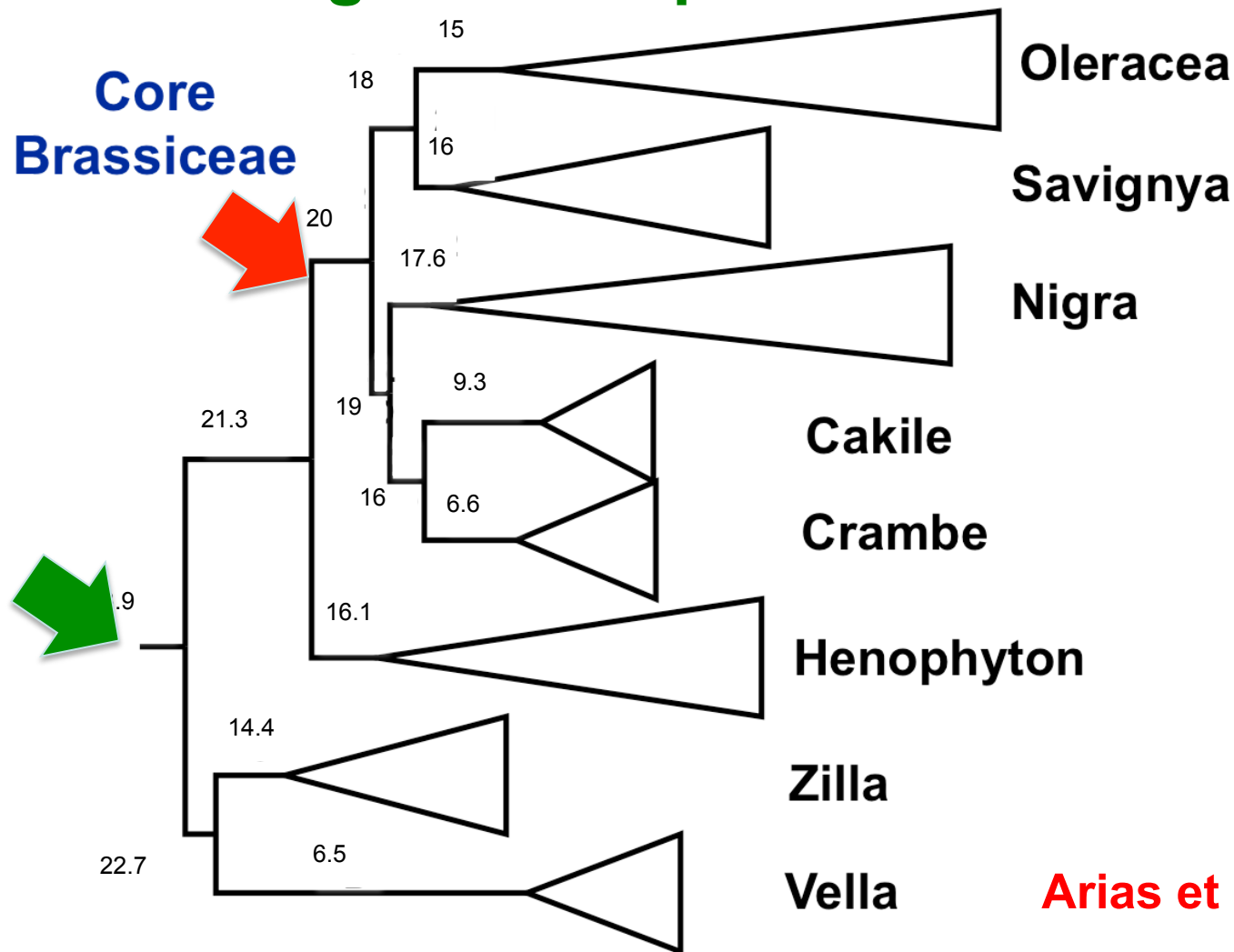


*Thlaspi* fossil  
30.8-29.2 Ma  
Beilstein et al. 2010

**Tribe Brassiceae colonized North Africa about 24 MYA and moved clockwise through Europe**



Diversification dates agree with previous results;  
however, there is lag time between the  
**whole genome triplication & diversification**



**Arias et al. 2014 AJB**

# Tribe Brassiceae Conclusions

**Arias et al. 2014. American Journal of Botany 101: 86 – 91**

**Arias and Pires. 2012. Taxon 61: 980 - 988**

- 1. A new phylogeny of the tribe Brassiceae led to discovery of a New African clade (Henophyton)**
- 2. Origin of tribe Brassiceae was inferred in the Saharo-Sindian region about 24 Ma. Ancestors colonized Europe through Western North Africa**
- 3. Diversification shifts in the tribe do not coincide with the whole triplication event**



# **Transcriptome Based Nuclear Tree Methods for Tribe Brassiceae**

- **RNA extracted from young leaf tissue**
- **Sequenced six species per lane on Illumina Hi-Seq with 2x100bp**
- ***De novo* transcriptome assembly and ortholog selection using the Agalma Pipeline (Dunn, et.al., 2013)**
- **Concatenated tree based on 1,663 loci with 72% occupancy estimated using RAxML (Stamatakis, 2014)**

# Outline: Phylogenomics

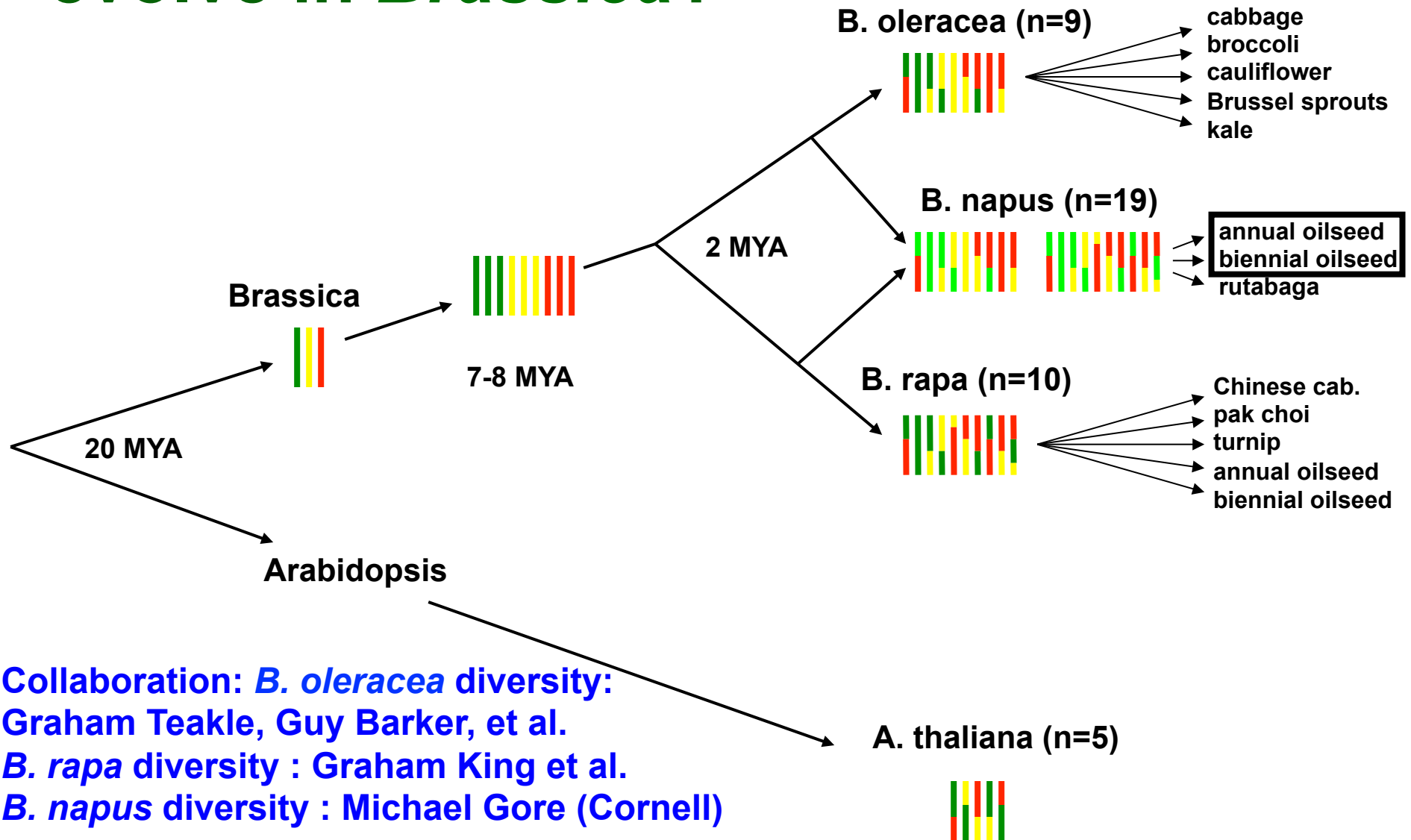
**I. Order Brassicales**

**II. Family Brassicaceae**

**III. Tribe Brassiceae**

**IV. Future Directions**

# How do genomes and networks evolve in *Brassica*?



Collaboration: *B. oleracea* diversity:  
 Graham Teakle, Guy Barker, et al.  
*B. rapa* diversity : Graham King et al.  
*B. napus* diversity : Michael Gore (Cornell)

# Acknowledgments - People

## MU Collaborators/Coauthors:

**Gavin Conant**

**Corey Hudson**

**Michael Bekaert**

**Pat Edger Zhiyong Xiong**

**Michelle Tang**

**Dustin Mayfield**



## External Collaborators/Coauthors:

**Eric Schranz Mike Barker**

**Marcus Koch Klaus Mummenhoff**

**Martin Lysak Mark Beilstein**

**Ihsan Al-Shehbaz**

**Team Butterfly: Chris Wheat, Heiko Vogel,  
Hanna Heidel-Fischer, David Heckel,  
Goran Hellekant**



# Acknowledgments - Funding

**NSF Plant Genome:**

**Functional Genomics of Plant Polyploids.**



**NSF Plant Genome Comparative Sequencing:**

**Toward unravelling the morphological plasticity and genome redundancy of *Brassica oleracea*.**

**NSF Tree of Life Program:**

**From Acorus to Zingiber: Assembling the Phylogeny of Monocots**

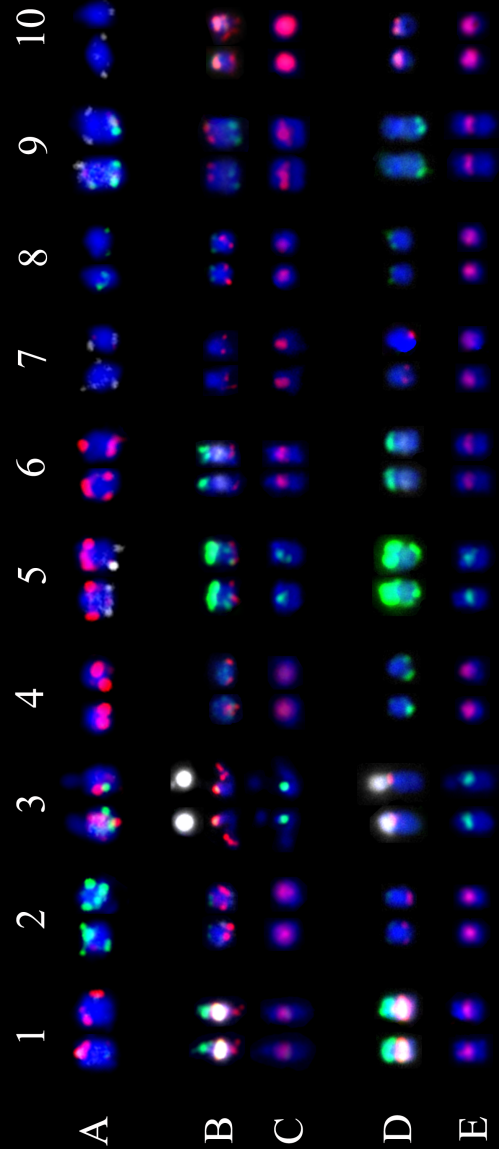
**MU Life Sciences Fellowships, LSUROP, UMEB**

**Society for Systematic Biology**

**Botanical Society of America**

**American Society of Plant Taxonomists**

# Thank You !



yellow flowers



albino stem



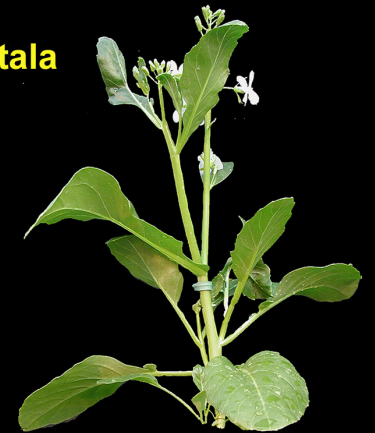
apetala



curly leaves



wild-type



glossy



yellow-green leaves



wrinkly leaves



dwarf