

University of Missouri
IPG Interdisciplinary Plant Group
Division of **Biological Sciences**
University of Missouri

Phylogenomics and Placement of Polyploidy with Transcriptomes in Brassicaceae

Dustin Mayfield
Wednesday, July 11th, 2012
University of Missouri-Columbia
J. Chris Pires Lab

Family History

- Three major lineages of ~3700 extant species
- Family-wide origin
 Pre-Miocene (molecular) or
 Miocene ~23-5MYA (paleobotanical/ecological)
- Rapid colonization of new dry environments
- Radiation and diversification rates among highest reported for angiosperms

Phylogenomics Transcriptomes in Brassicaceae

RNA (normalized) for phylogenomic resolution with 100s-1000s of markers is cost effective

Leveraging skewed transcript abundance by RNA-Seq to increase the genomic depth of the tree of life

Chris Todd Hittinger^{1,2}, Mark Johnston^{1,2}, John T. Tossberg^{1,2}, and Antonis Rokas^{1,2}

NATURE | LETTER

Michael R. McKain¹, Norman Wickett^{1,2}, Saravanaraj Aravamudan¹, W. Richard Mark W. Chase^{1,2}, J. Chris Pires¹, Claude W. Leebens-Mack^{1,2}

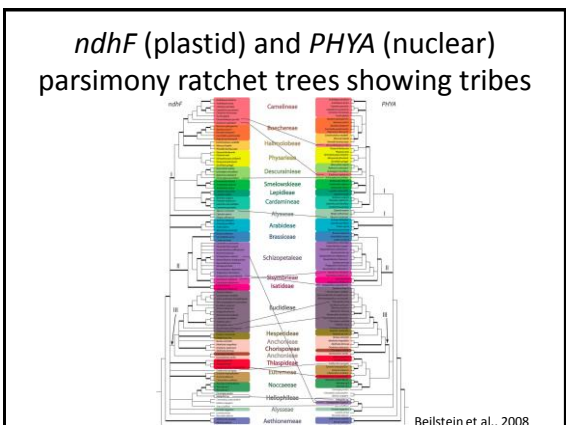
Phylogenomics reveals deep molluscan relationships

Kevin M. Kocot, Johanna T. Cannon, Christiane Todt, Mathew R. Citarella, Andrea B. Kohn, Achim Meyer, Scott R. Santos, Christoffer Schander, Leonid L. Moroz, Bernhard Lieb & Kenneth M. Halanych

Report

Transcriptomes of the Parasitic Plant Family Orobanchaceae Reveal Surprising Conservation of Chlorophyll Synthesis

Norman J. Wickett^{1,2}, Loren A. Honas¹, Eric K. Wafar¹, Malay Das^{1,2}, Kan Huang³, Biao Wu⁴, Lena Lanzher¹, Michael P. Timko¹, John Yoder¹, James H. Westwood², Claude W. Leebens-Mack^{1,2}



Method

- 1: Make libraries
- 2: Sequence
- 3: Assemble transcripts
- 4: Identify orthologs
- 5: Align
- 6: Make trees and networks
- 7: Placement of polyploidy

Method

- 1: Make libraries

For RNA isolation we use PureLink RNA minikits
For RNA-seq library preparation, we use TruSeq RNA Sample Preparation Kits v2

What have we learned?

Wet lab is easy, assembly/analysis is harder, typos, RNA quantity, Brassica vs Grasses, **Quality RNA!**

Workflows for Illumina library preparation (RNA and DNA) and suggestions soon @ <http://dustinmayfieldjones.wordpress.com/> under the 'For Collaborator's tab

Dustin Mayfield-Jones
 Graduate Student at the University of Missouri-Columbia

Method

- 2: Sequence – ~190M, 170M filtered

Method

- 3: Assemble transcripts - Trinity

Method

- 4: Identify orthologs

Identified set of 959 single copy genes that are shared single copy genes for *Arabidopsis thaliana*, *Populus trichocarpa*, *Vitis vinifera*, and *Oryza sativa*, or APVO SSC genes

Taxonomic Group	Number of single copy APVO SSC Plant Tribes present
Eurosids	913
Asterids	519
Core Eudicots	76
Basal Eudicots	189
Monocots	948
Basal Angiosperms	48
Gymnosperms	502
Vascular Plants	438
Green Algae	190

DFG **HaMStR**
 (Hidden Markov Model based Search for Orthologs using Reciprocity)

Modified from Duarte et al., 2010

Method

- 5: Align – MUSCLE

MUSCLE: multiple sequence alignment with high accuracy and high throughput by Robert C. Edgar

Alternatives include ClustalW, T-coffee, or other multiple sequence aligner.

Method

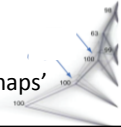
- 6: Make trees and networks – RaxML / PAUPRat

Dunn et al., 2008

Trees vs Networks

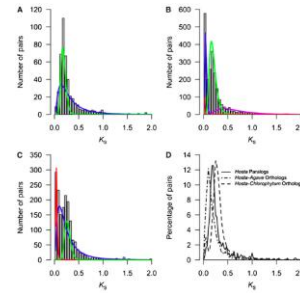
- Unlike a phylogenetic tree, which shows the most likely or most parsimonious tree, a network shows many (or all) feasible trees in one 2D representation. In effect, it allows us to visually represent the ambiguity we have about which of the feasible trees is the true tree.

I dream of beautiful trees as 'heat maps'



Method

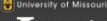
- 7: Placement of Polyploidy – Ks plots




McKain et al., 2012

Summary


- Use RNA-seq to help **resolve the complex phylogenomic and WGD history** of Brassicaceae, a family with high radiation and diversification rates, multiple rounds of WGD, and hybridization.
- Inform those interested in the evolutionary relationship of both the **model plant, Arabidopsis, and crops, such as the Brassica oilseed and vegetable crops** including agronomically important cultivars such as canola, broccoli, cauliflower, cabbage, kale, Kohlrabi, and Brussels sprouts
- We expect to **identify previously unknown WGD events** that will further our understanding about their genetic consequences.



University of Missouri




Interdisciplinary Plant Group




Division of Biological Sciences

University of Missouri




Pires Lab –
Chris Pires, Pat Edger, Tatiana Arias, Madison Knapp, Avinash Karn




Marcus Koch, University of Heidelberg
Martin Lysák, Masaryk University

Funding and support:
University of Missouri Life Science Fellowship
National Institutes of Health T32 training grant



NATIONAL INSTITUTES OF HEALTH



Transforming life science through
COLLABORATION × COMPUTATION × COMMUNICATION

A HOWARD HUGHES MEDICAL INSTITUTE PROGRAM AT MIZZOU