USING EVOLUTIONARY RATE COVARIATION TO IDENTIFY CANDIDATE GENES IN PLANTS

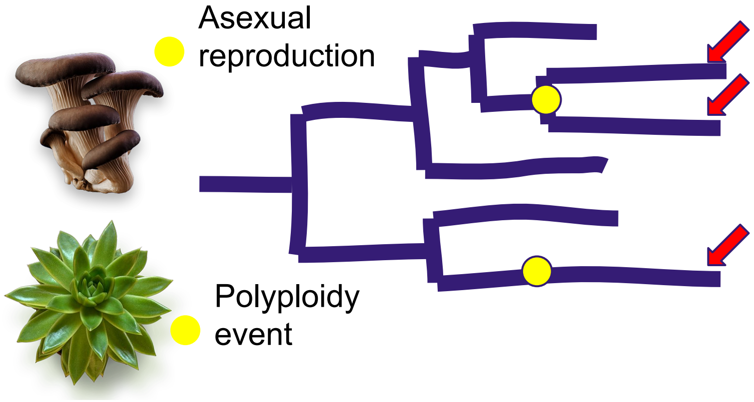
Colby Witherup

Dr. Norman Wickett and Dr. Matthew Johnson

Chicago Botanic Garden

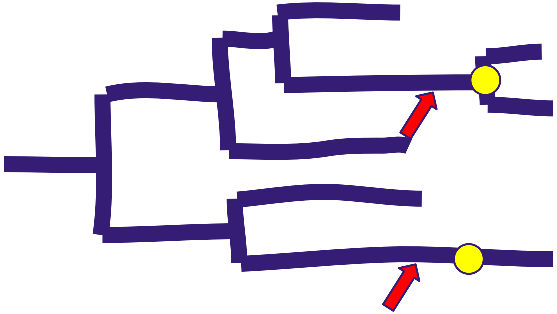
Northwestern University

Some genes evolve faster **following** the emergence of a trait.



Some genes may evolve faster **preceding** the emergence of a trait.

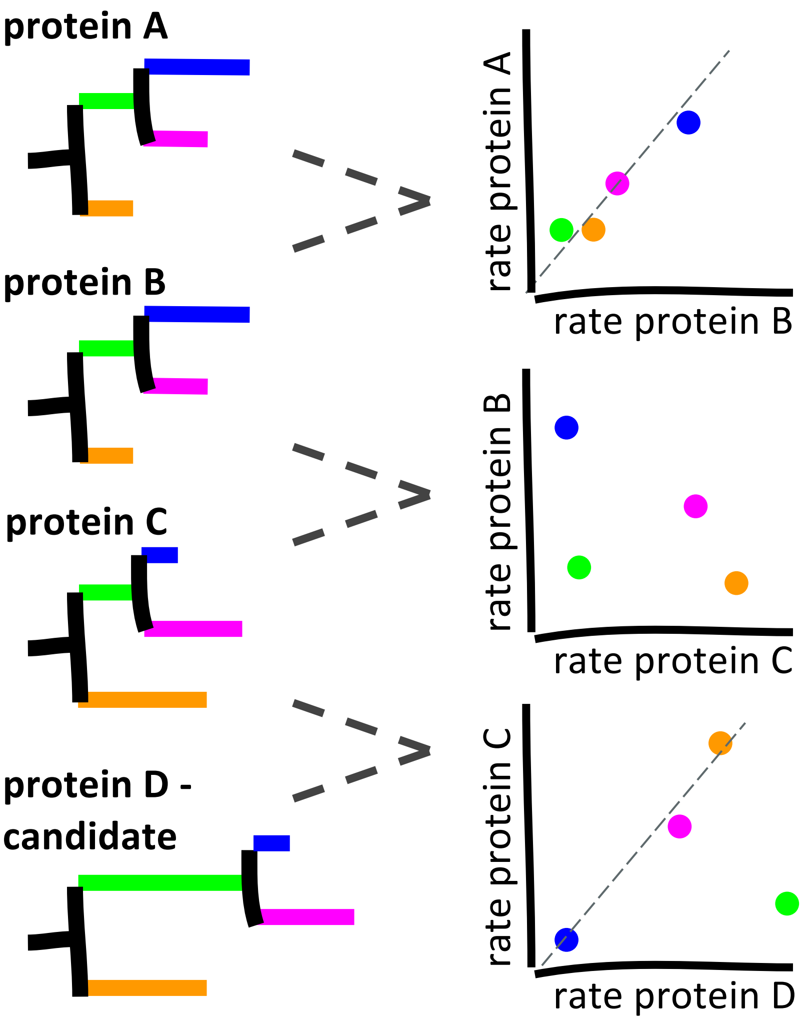
More difficult to study



Evolutionary Rate Covariation

Developed by Dr. Nathan Clark's lab at the University of Pittsburgh (2012)

Most genes that are involved in the same function evolve slowly and as a group.



A and B have high ERC and are likely involved in the same function

B and C have low ERC and are likely not involved in the same function

C and D have correlated evolutionary rates **except** for on one branch of the phylogeny

COPPERPOT

pipeline

My bioinformatic pipeline for calculating ERC in plants: Covariation Of Plant Protein Evolutionary Rates with Pruning Of Trees

“The computing resources needed to handle genome data will soon exceed those of Twitter and YouTube.” - Nature News July 2015

Genomes

Orthologs HPC

Pruning UNIQUE TO PLANTS

Alignments HPC

Trees HPC

ERC Results

A single human genome is ~90 gigabytes. Plant genomes can be both smaller and larger.

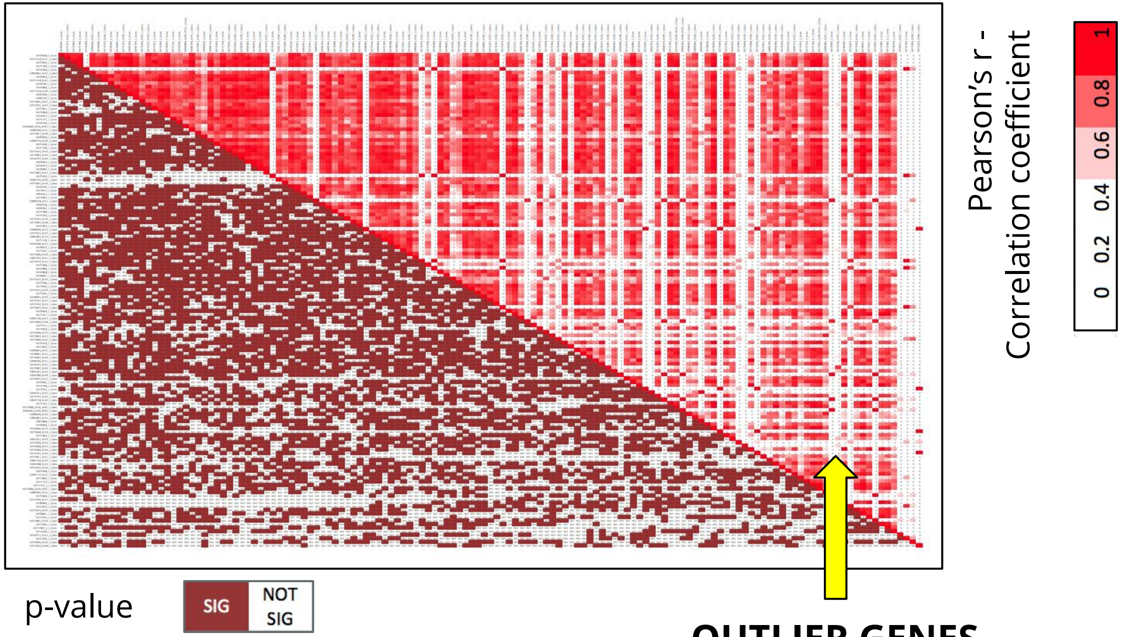
Each genome is compared to each other genome to find matching genes. (N-squared pairwise comparisons.

Plants have high levels of gene duplication, but ERC requires only a single copy of each gene. COPPERPOT allows users to designate duplication nodes on gene trees.

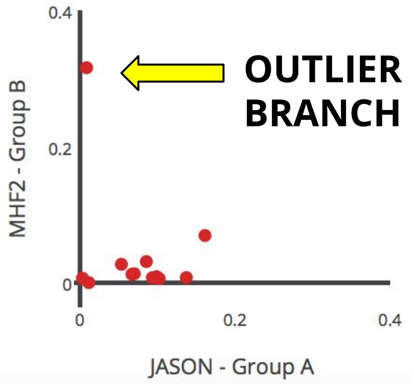
PRELIMINARY RESULTS

The pipeline is able to identify individual genes with unique rates of evolution on individual branches of a phylogeny.

Example from the turnip family:
I measured ERC on ~300 meiosis genes across 9 species. One group of 7 genes shows strong correlation with each other, but not with other meiosis genes. The scatterplots show an outlier branch - a branch that corresponds to an interesting meiotic transition.



OUTLIER GENES



OUTLIER BRANCH

Capsella grandiflora
Outcrossing

Capsella rubella
Selfing

Next steps

I'll be using my pipeline to investigate meiosis genes in three unrelated plant genera that each show a transition between strict diploidy and a tendency toward polyploidy (Whole Genome Duplication).

A. An ideal study clade
B. *Malus* (apple)
C. *Castilleja*
D. *Fragaria* (strawberry)