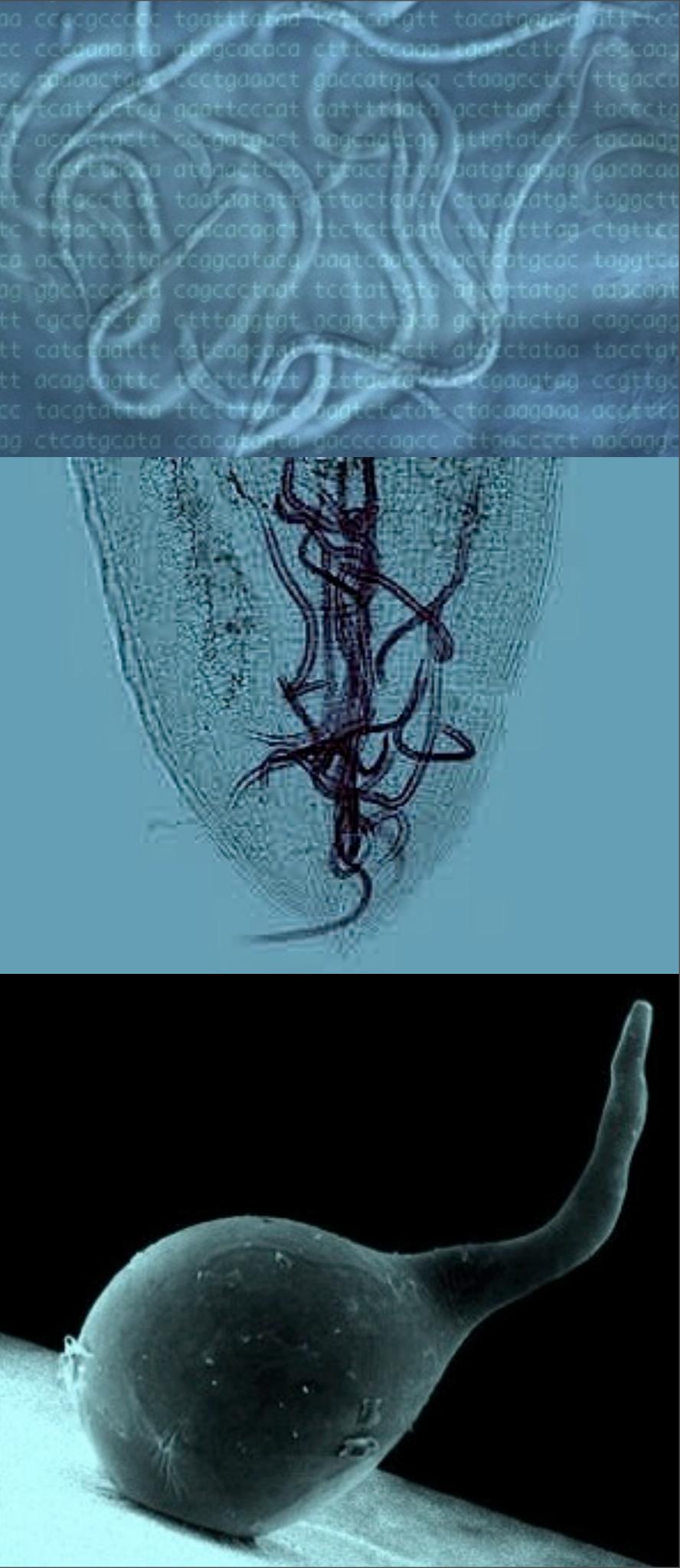


# THE COMPARATIVE GENOMICS OF 'ANCIENT ASEXUALITY' AND HYBRIDIZATION IN ROOT KNOT NEMATODES

Dave Lunt  
Evolutionary Biology Group, University of Hull



# COMPLEX HYBRID ORIGINS OF ROOT KNOT NEMATODES

Dave Lunt

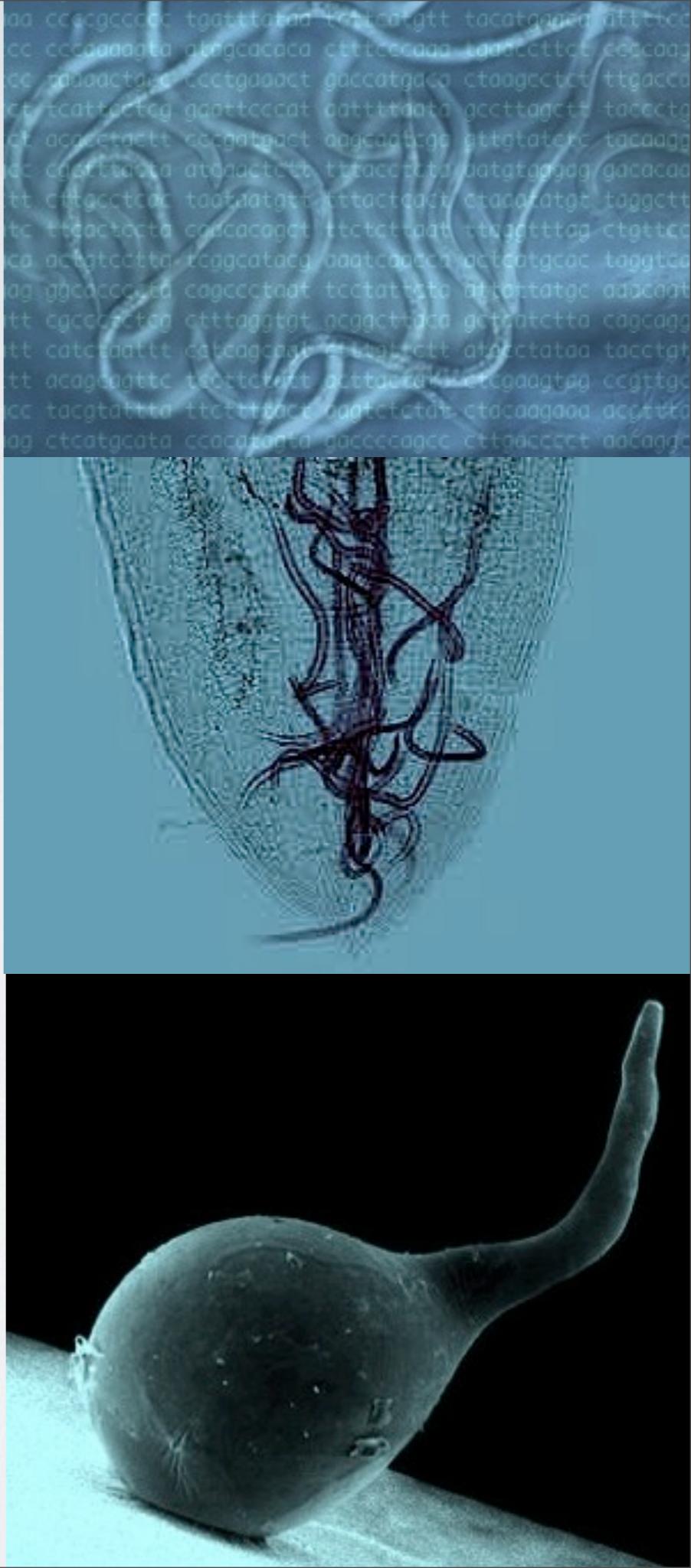
Evolutionary Biology Group, University of Hull

Sujai Kumar

Georgios Koutsovoulos

Mark Blaxter

Institute of Evolutionary Biology, University of Edinburgh



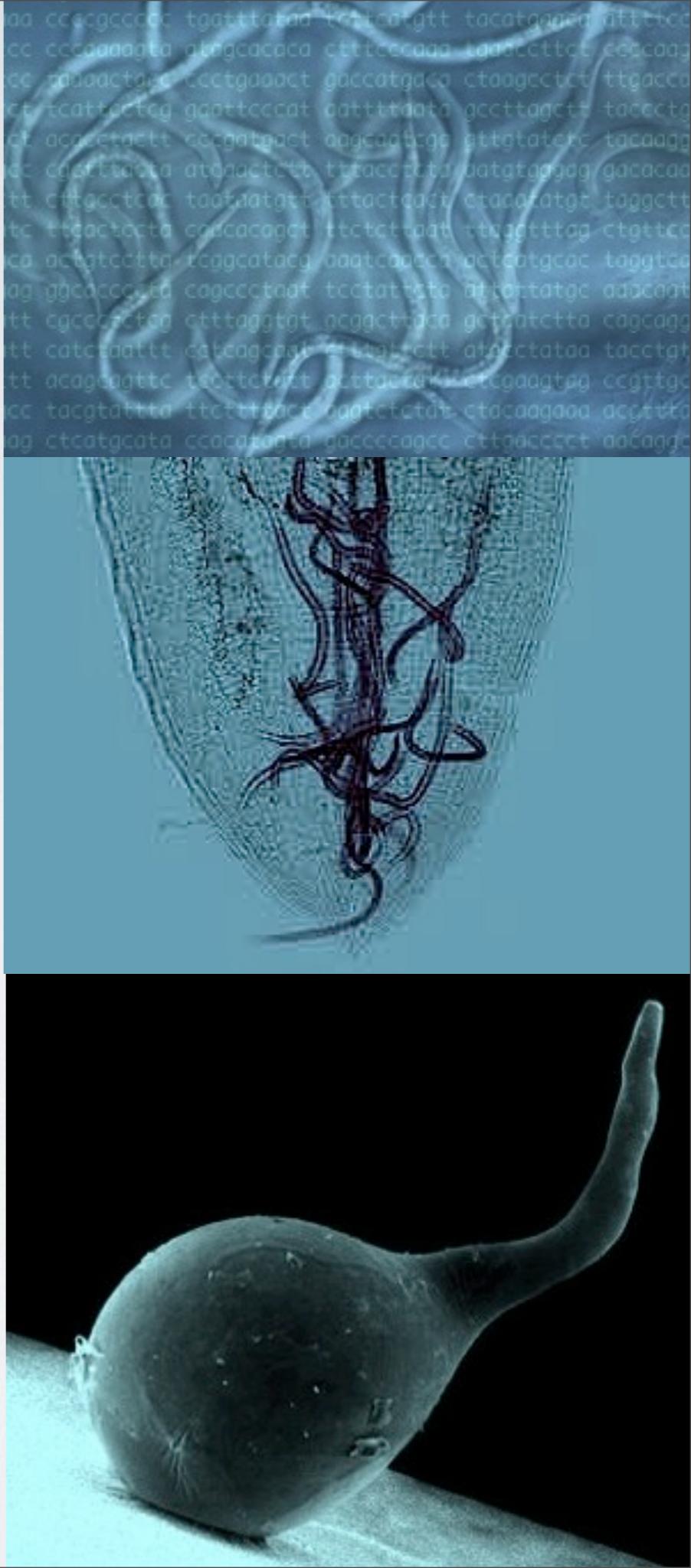
# COMPLEX HYBRID ORIGINS OF ROOT KNOT NEMATODES

## Acknowledgements

Africa Gómez, Richard Ennos, Amir Szitenberg,  
Karim Gharbi, Chris Mitchell, Steve Moss, Tom  
Powers, Janete Brito, Etienne Danchin, Marian  
Thomson & GenePool

## Funding

NERC, BBSRC, Yorkshire Agricultural Society,  
Nuffield Foundation, University of Hull,  
University of Edinburgh



# COMPLEX HYBRID ORIGINS OF ROOT KNOT NEMATODES

## Dave Lunt

Evolutionary Biology Group, University of Hull



@EvoHull



+EvoHull



davelunt.net



dave.lunt@gmail.com



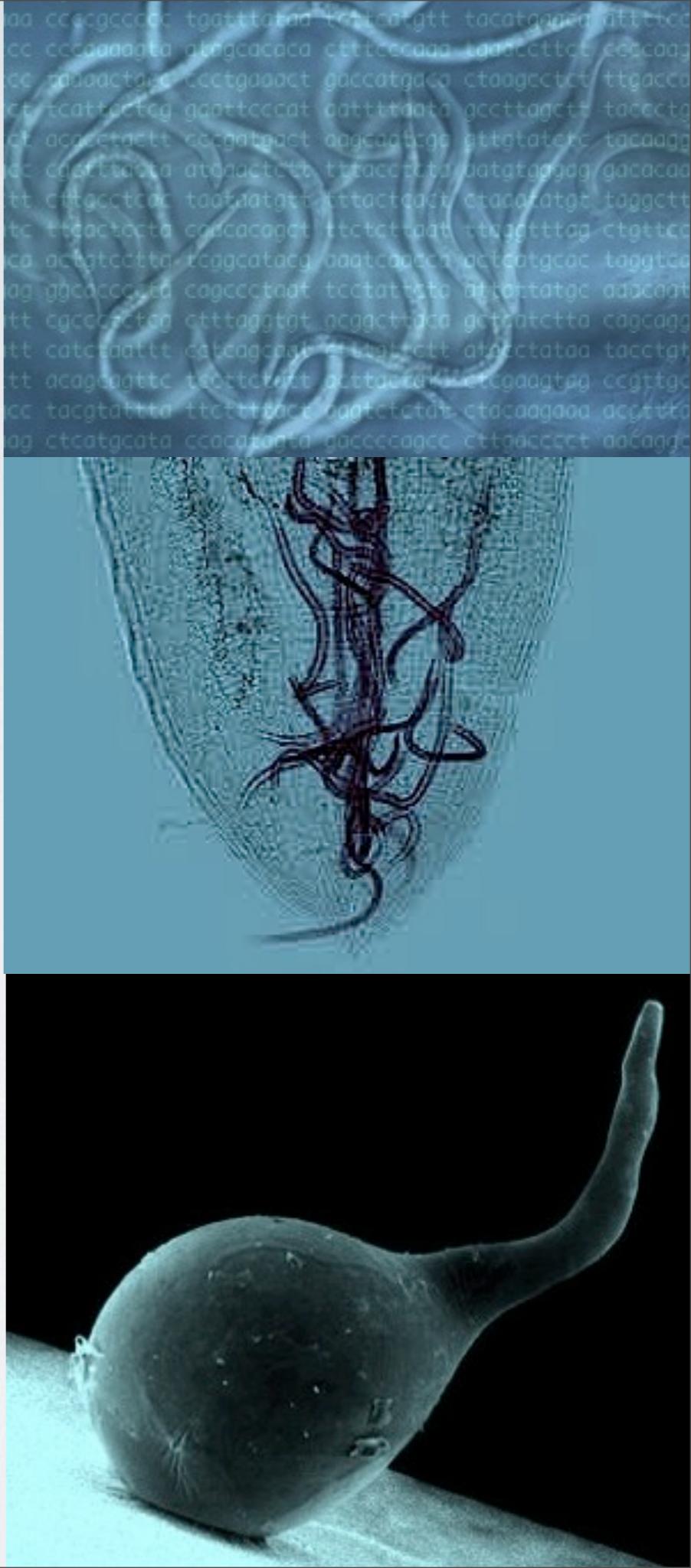
@davelunt



+davelunt



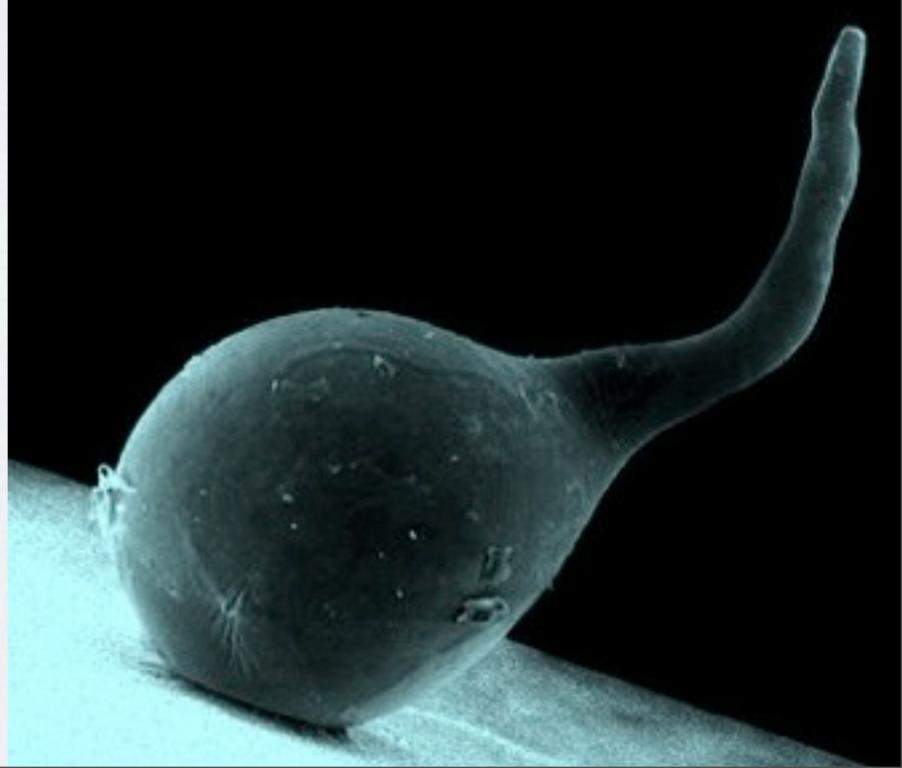
<http://www.github.com/davelunt>



COMPLEX HYBRID ORIGINS OF ROOT KNOT NEMATODES

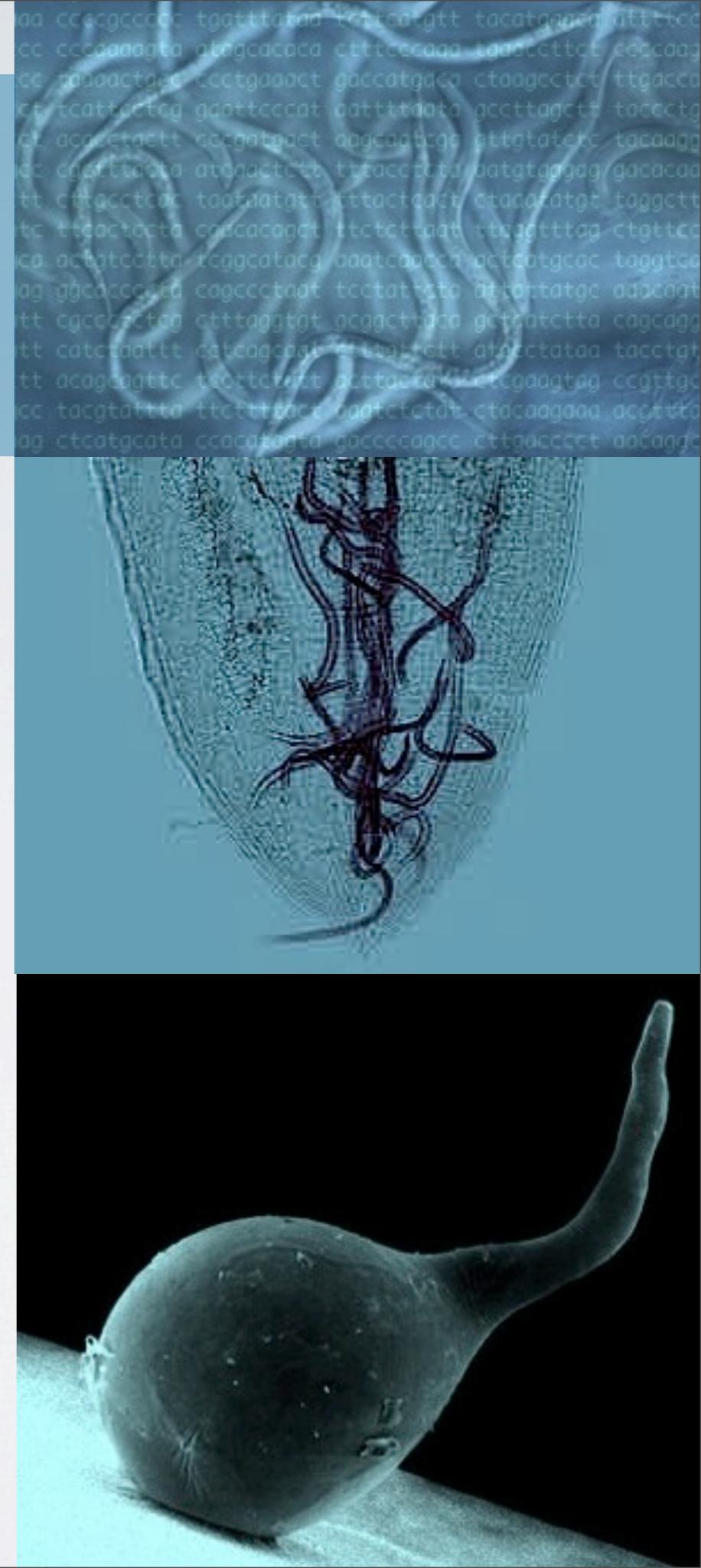
# WHAT'S IN A GENOME & WHY?

aa ccccgcccccc tgattttatcc ttatgttt tacatgggg attttcc  
cc cccggggta gtcgcacaca ctttcccaaa tttttttttt ccccaag  
cc ggaaactggcc ccctgaaact gaccatgaca ctaagccctt ttgaccc  
ct tcattttcg gatttcccat cattttataa gccttagctt tacactg  
ct acacacttttccggatact aagcaatcgg atttgtatctc tacaagg  
tc cgtttttttt ataaaaatcc ttatccctttt atgtggggdg gacacaa  
tt ctgtccctcc taataatgtt ttatctactt atacatatgt taggttt  
tc ttcaactcttcc cccacacat ttctctttttt ttcgggtttt tagttcc  
ca actatcccttcc cgccatccatg aaatccatcc tttcatgcac taggttt  
ag ggcccccata cagcccttatc ttctatccat ttatccatgc adccat  
tt cggccatctt ctttaggtgtt cccggcttccat gctatcttcc cagcagg  
tt catccatccat cttccatccat ttatccatgc tttccatccat tttccatcc  
tt acagccatccat ttatccatccat ttatccatgc tttccatccat tttccatcc  
cc tacgtattttcc ttatccatccat ttatccatgc tttccatccat tttccatcc  
og ctcatgcata ccacatccat ccccccaccc ctttccatccat tttccatcc



# WHAT'S IN A GENOME & WHY?

For many eukaryotes:  
mostly transposons,  
repeats, & sequences  
of *incertae sedis*  
  
but why?



# WHAT'S IN A GENOME & WHY?

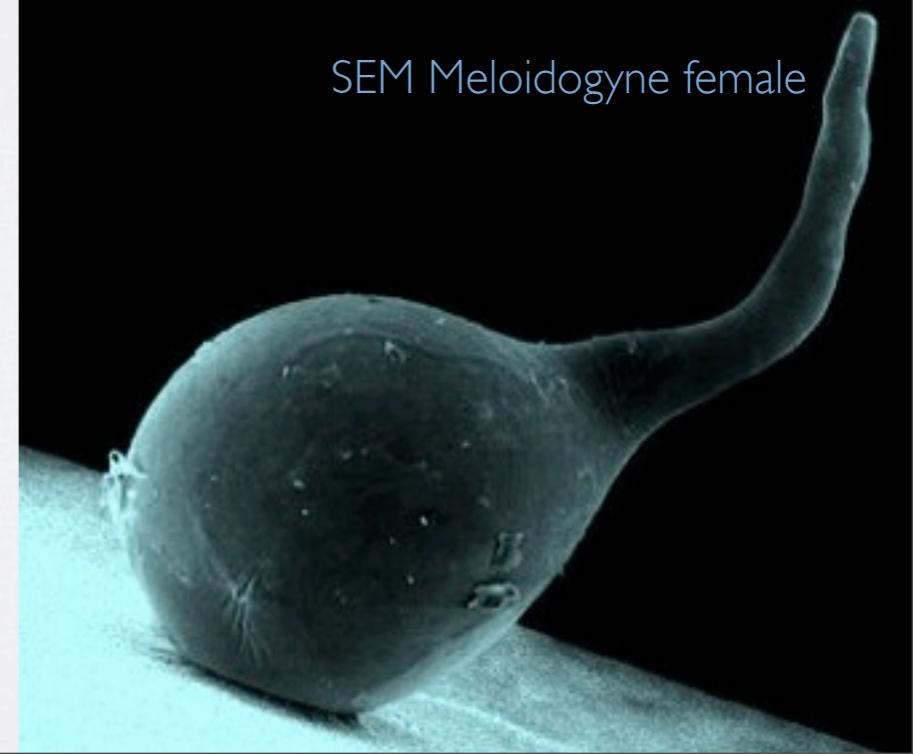
## Evolutionary Forces:

- Selection
- Gene Flow
- Mutation
- Drift
- Recombination

aa ccccgcccccc tgatttataa tttatgtt tacatggggattttcc  
cc cccggatgtt gtcgcacaca cttcccaaa tttttttttt ccccaag  
cc ggaaacttcc ccctgaaact gaccatgaca ctaagccctt ttgaccc  
ct tcatttcgt gatttcccat cattttataa gccttagctt tacatgt  
ct acatctactt ccggatactt agggatcggg tttgtatctc tacaagg  
tc cgtttataa ataaactttt ttatccctt atatgtggdg gacacaa  
tt cttgcctcac taataatgtt ttatctactt ctatatatgt taggttt  
tc ttcaactcttcc cacacatgtt ttctttttt atgggttttag ctgttcc  
ca actatcttcc tcggcatacg aaatccatcc tttatgtac taggttt  
ag ggcccccata cagcccttat tccatttttcc ttatattatgc adccat  
tt cggccatcg ctttaggtgtt cccgcttccaa gctatcttcc cagcagg  
tt catccattt ccgcggccat tttttttt atatctataa tacctgt  
tt acageggttc tccctctttt tttatccatcc tccaaatggccatcc  
cc tacgtatttt ttcttttccat tttatctatcc ctacaagaaaa acccttcc  
ag ctcatgtatccat ccacatgttcc gccccccatcc ctttcccttccatcc



SEM Meloidogyne female



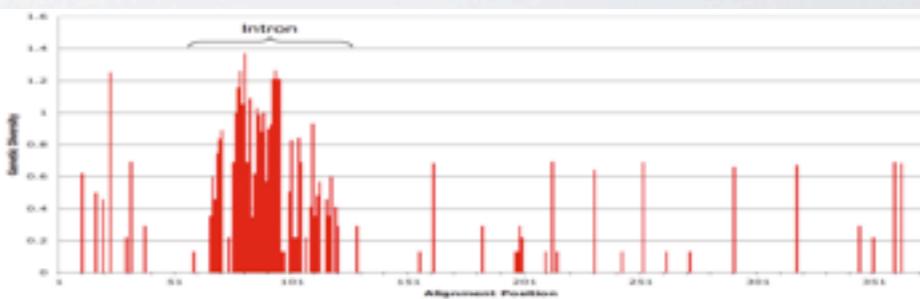
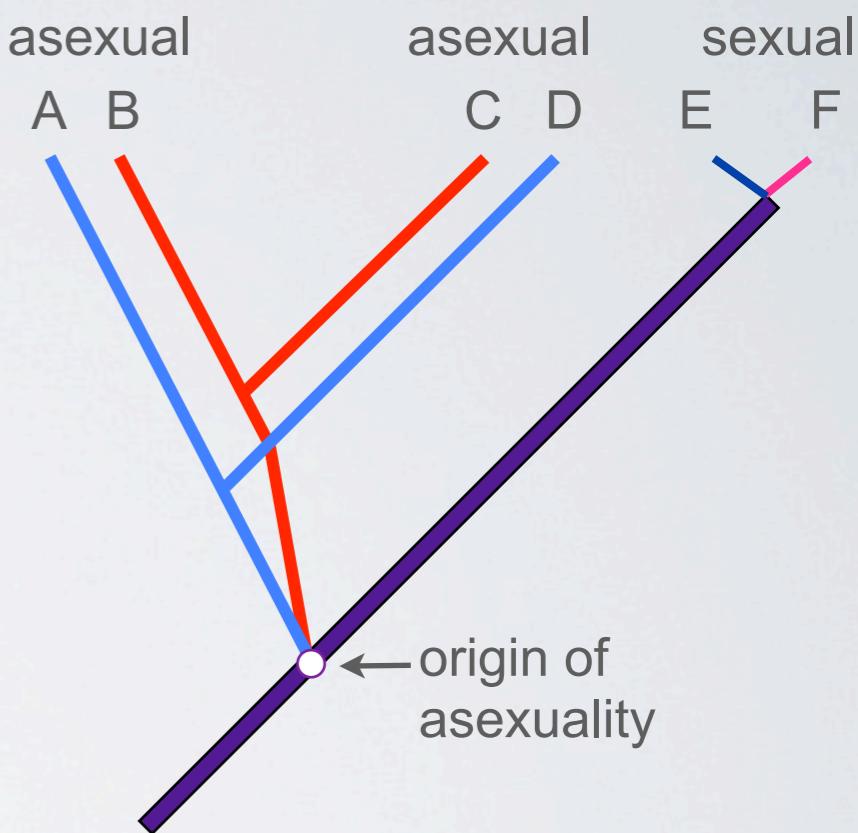
# Studying Recombination

Study its effects in genomic regions with reduced recombination

- sex chromosomes
  - inversions

Study its action in species that have lost meiotic recombination

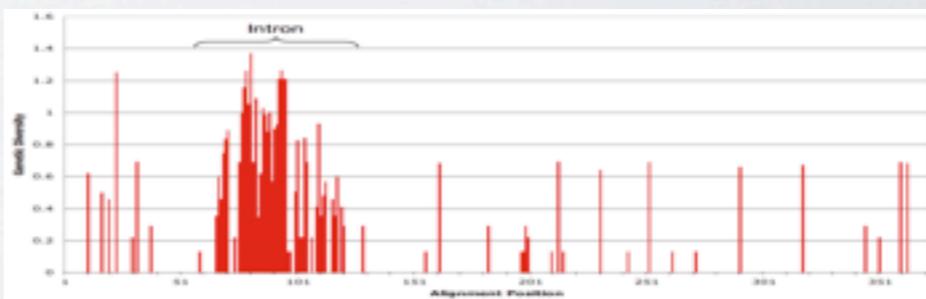
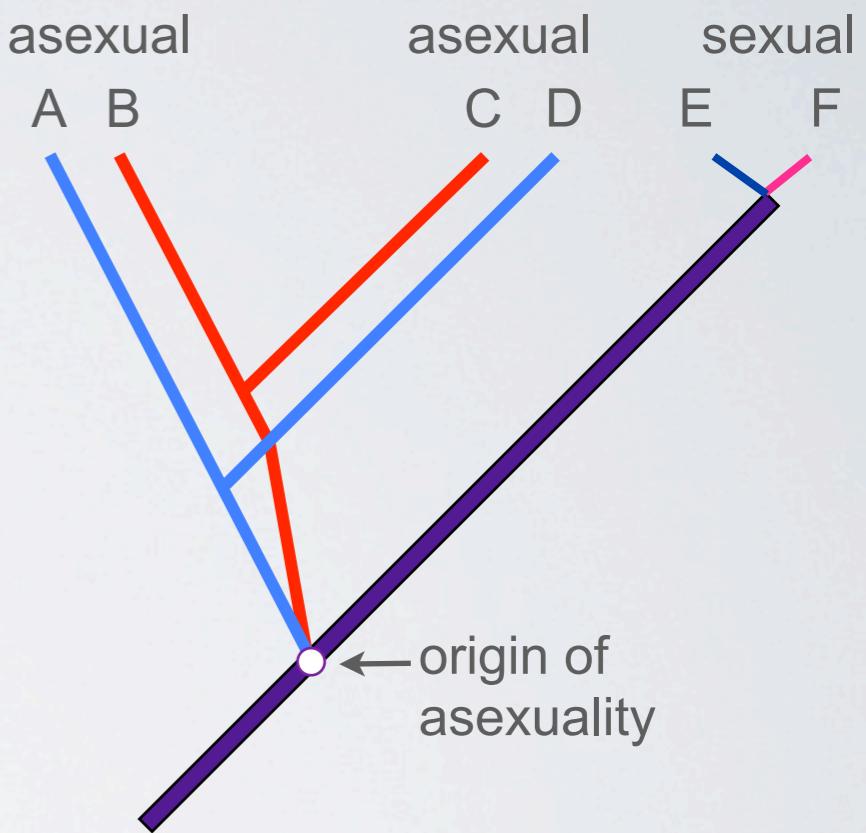
- asexuals



# Studying Recombination

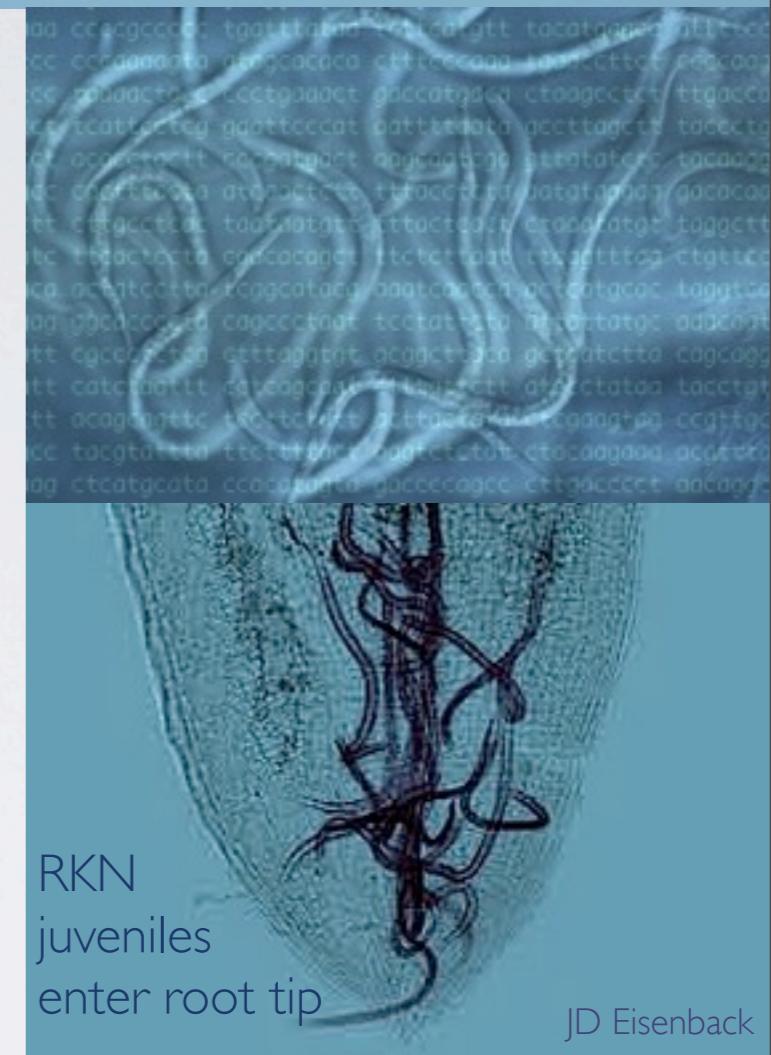
# Mitotic reproduction has consequences for the genome

- decay of sex-specific genes
  - extreme Allelic Sequence Divergence
  - loss of mutational effects of recombination



# Meloidogyne Root Knot Nematodes

- Globally important agricultural pest species
- Enormous plant host range
  - parasitize all main crop plants
  - ~5% loss of world agriculture



JD Eisenback

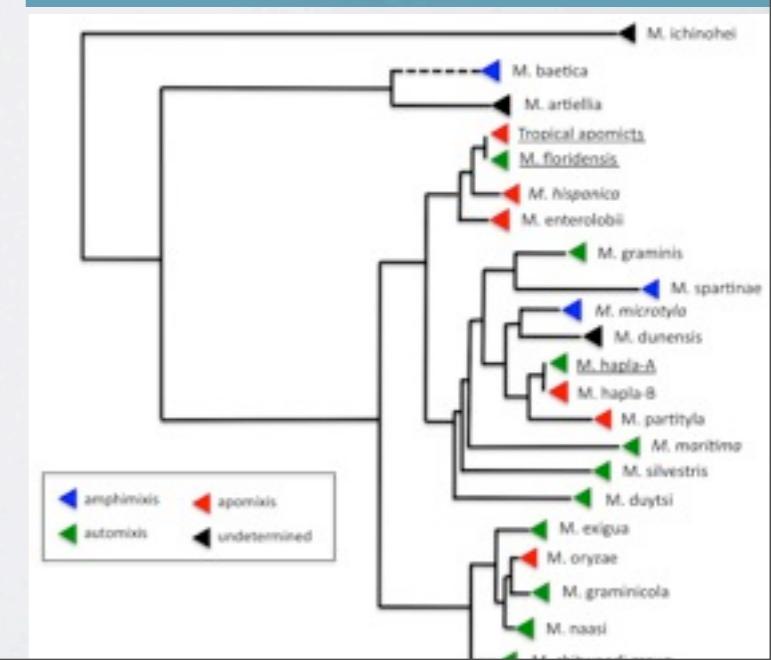
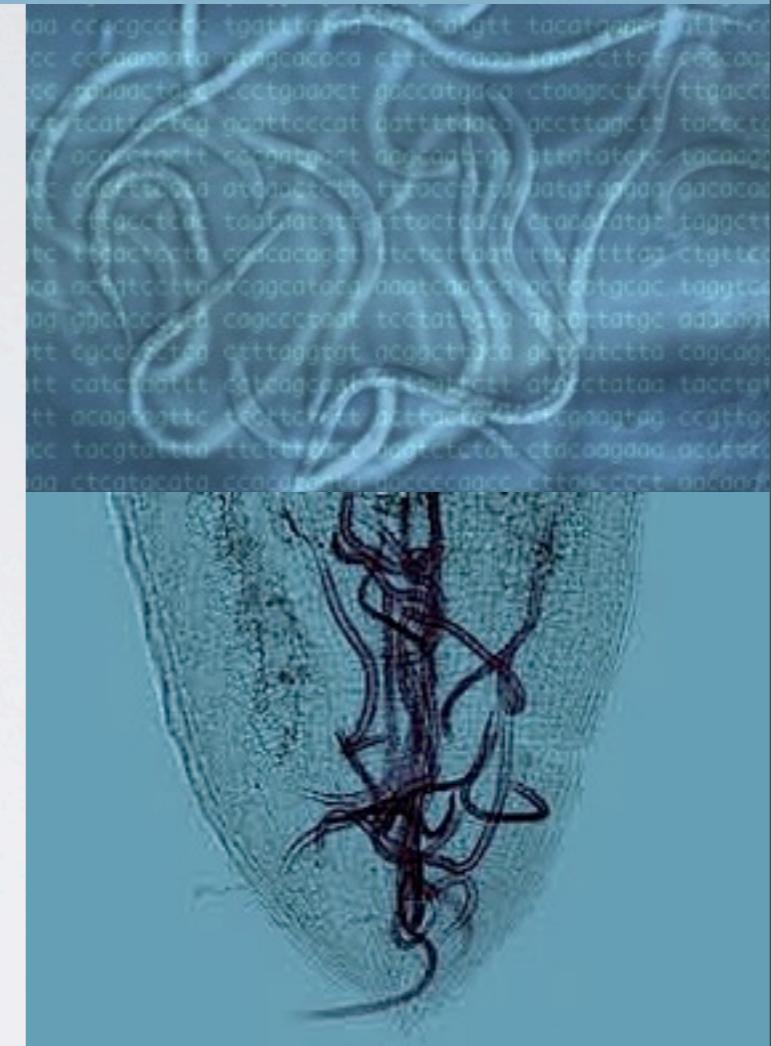
SEM Meloidogyne female



JD Eisenback

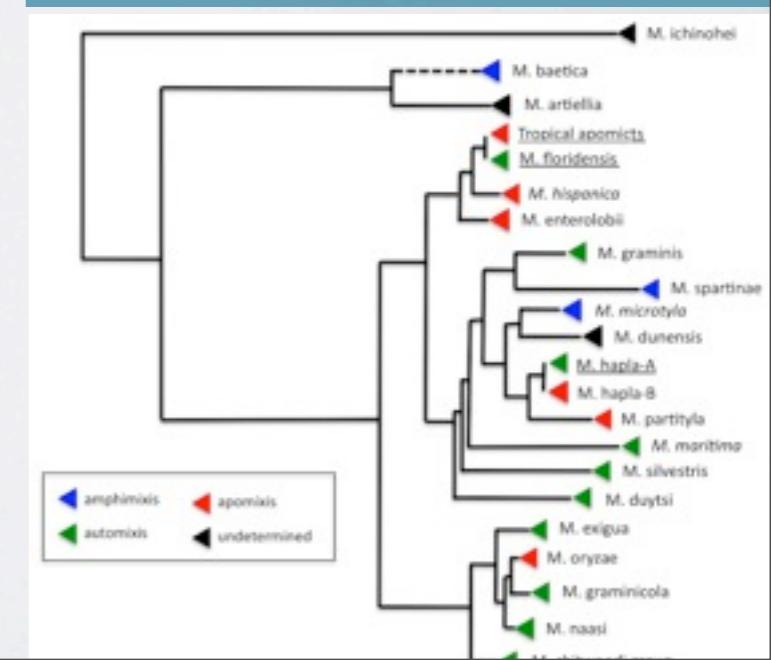
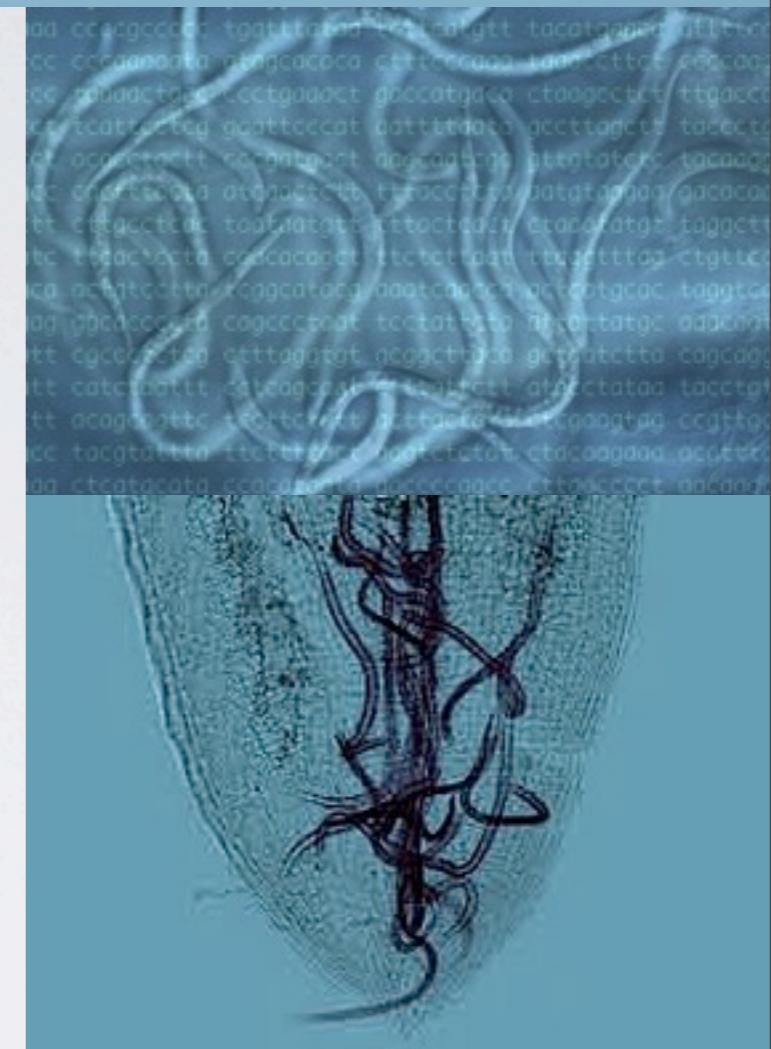
# Meloidogyne Reproduction

- Wide variety of reproductive modes in a single genus
- Many species are mitotic parthenogens without chromosome pairs
  - Incapable of meiosis
  - Could be ‘ancient’ asexuals
    - 17 million years without meiosis?



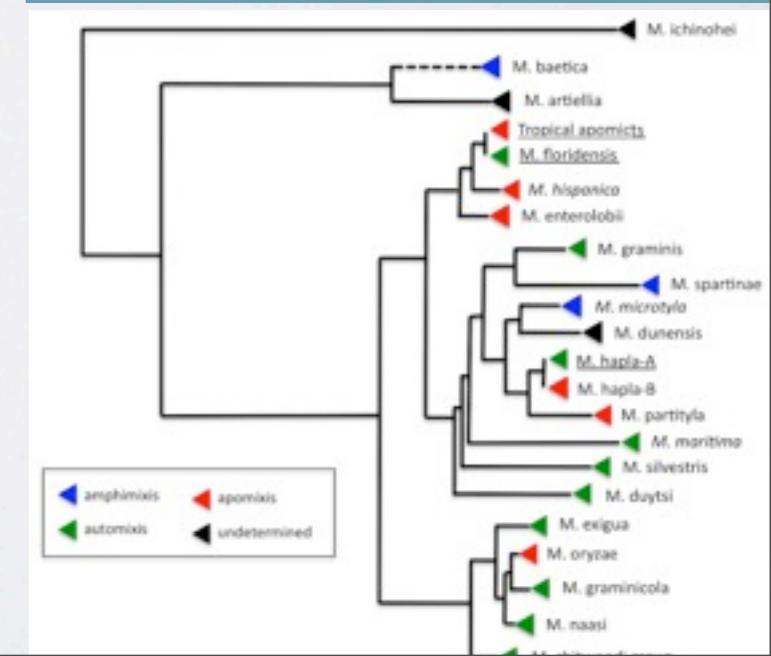
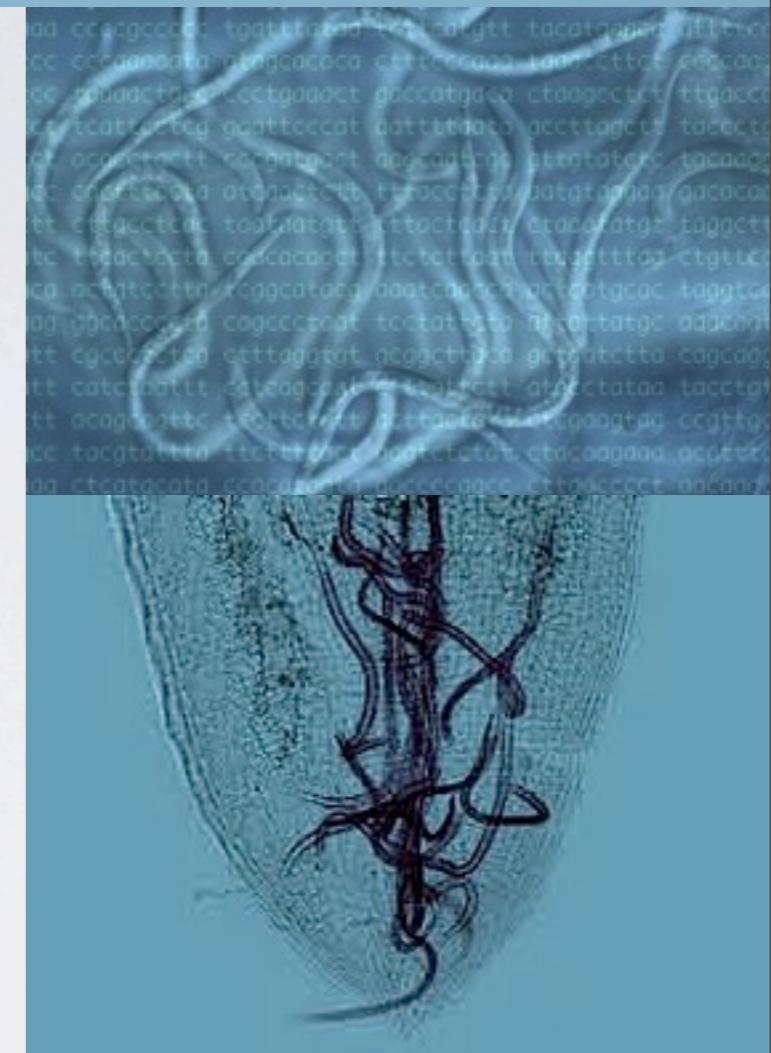
# Meloidogyne Reproduction

- Wide variety of reproductive modes in a single genus
- Many species are mitotic parthenogens without chromosome pairs
- Other species are meiotic parthenogens
  - automixis



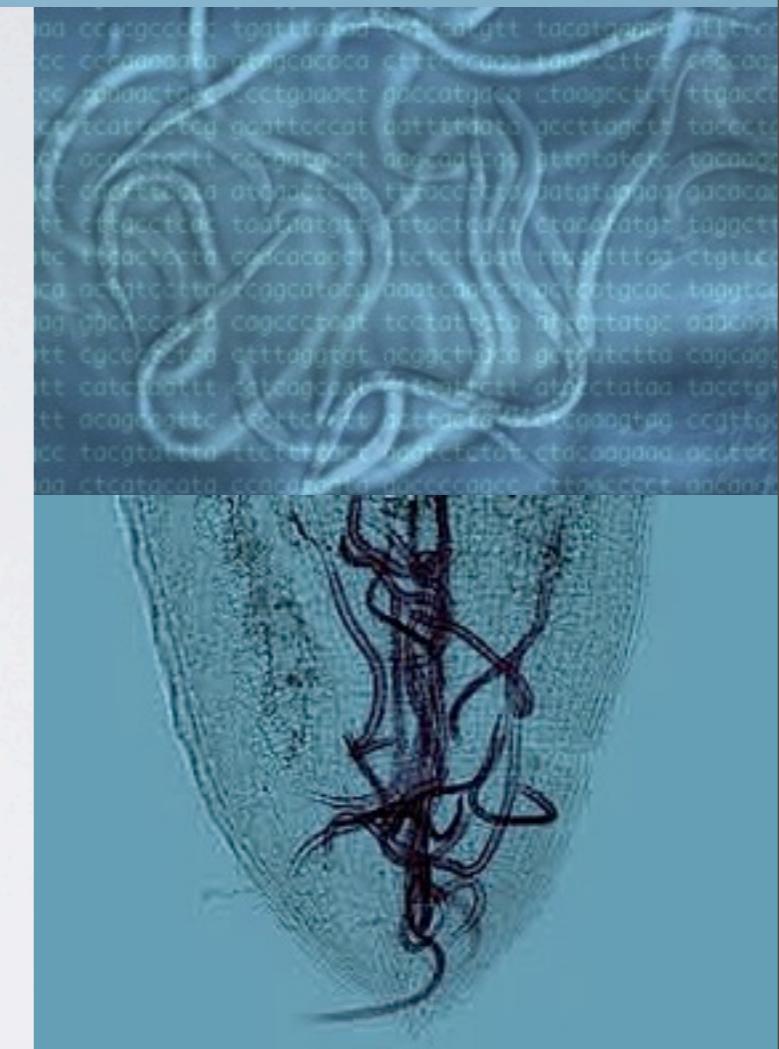
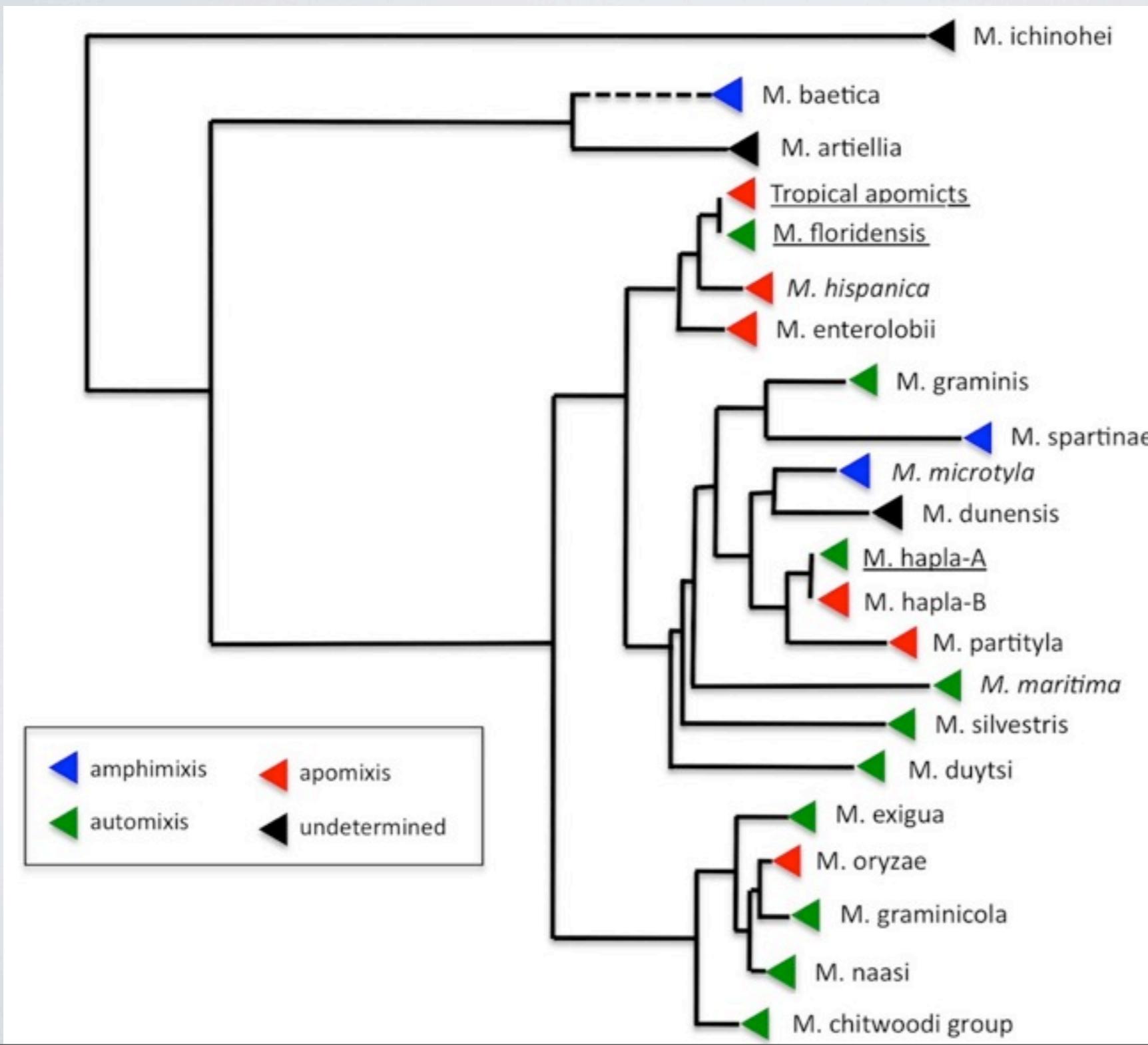
# Meloidogyne Reproduction

- Wide variety of reproductive modes in a single genus
- Many species are mitotic parthenogens without chromosome pairs
- Other species are meiotic parthenogens
- Some species are obligatory outbreeding sexuals with males & females
  - amphimixis



# Meloidogyne Reproduction

- Wide variety of reproductive modes in a single genus

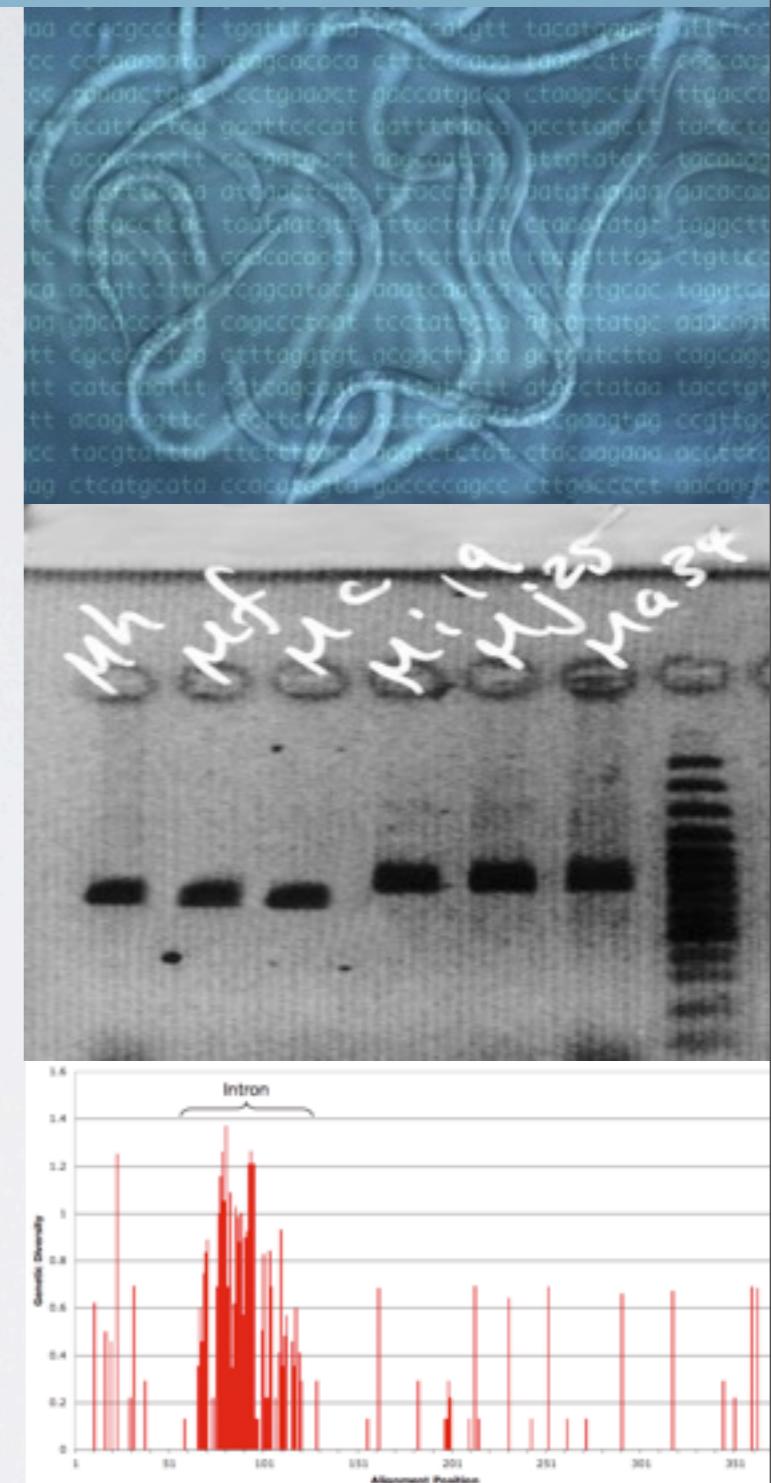


# Are RKN Ancient Asexuals?

Lunt DH 2008 BMC Evolutionary Biology 8:194

Investigations based on multi-species single gene sequencing and phylogenetics:

- testing for extreme allelic sequence divergence
- testing for changes in molecular evolution pattern of sex-specific loci

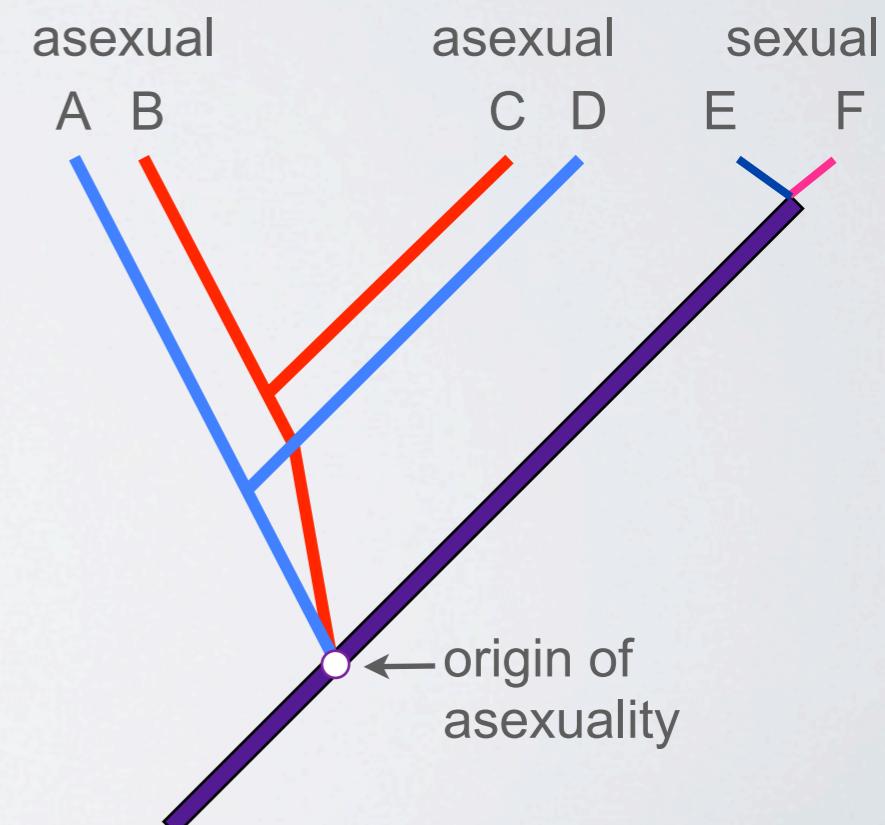


# Extreme Allelic Sequence Divergence

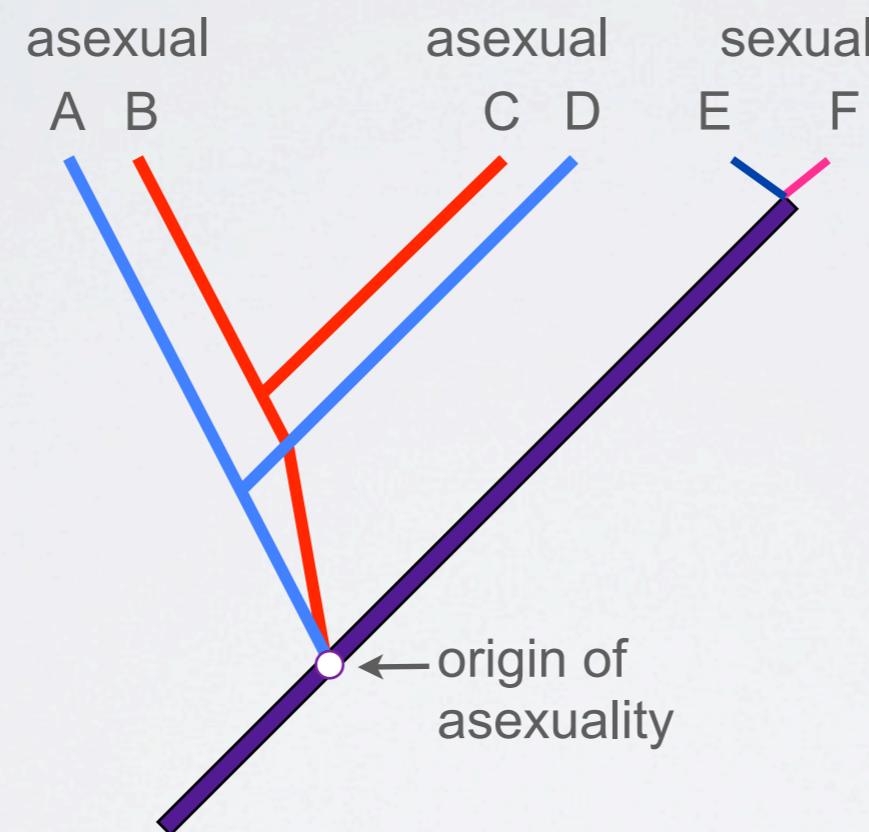
MJD White 'Animal Cytology and Evolution' 1<sup>st</sup> ed 1945, p283

"If we suppose an ameiotic form evolving for a very long period of time we might imagine its two chromosome sets becoming completely unlike, so that it could no longer be considered as a diploid either in a genetical or cytological sense."

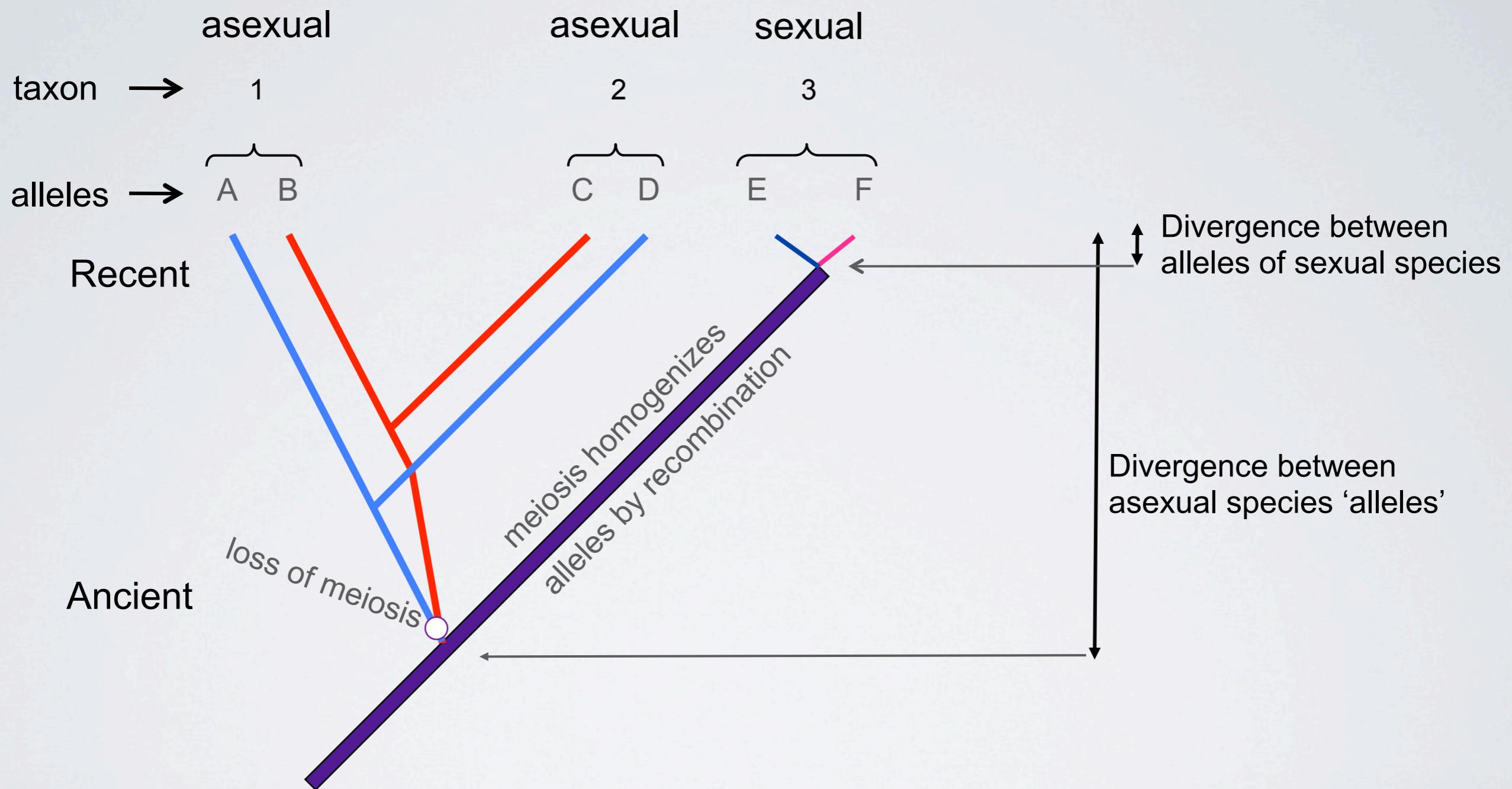
Sometimes called Meselson effect, similar to paralogous loci



# Extreme Allelic Sequence Divergence

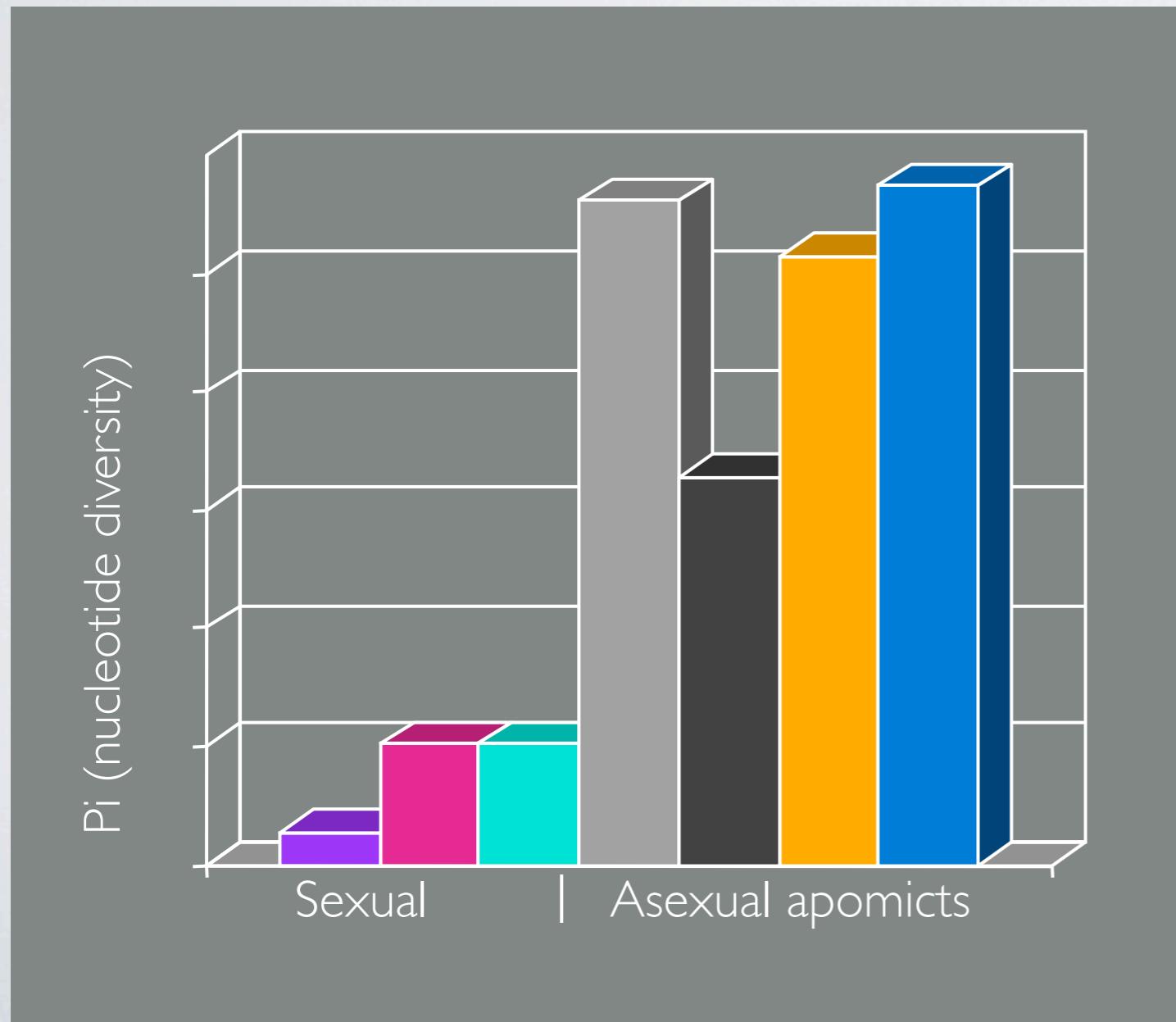


# Extreme Asexual ASD

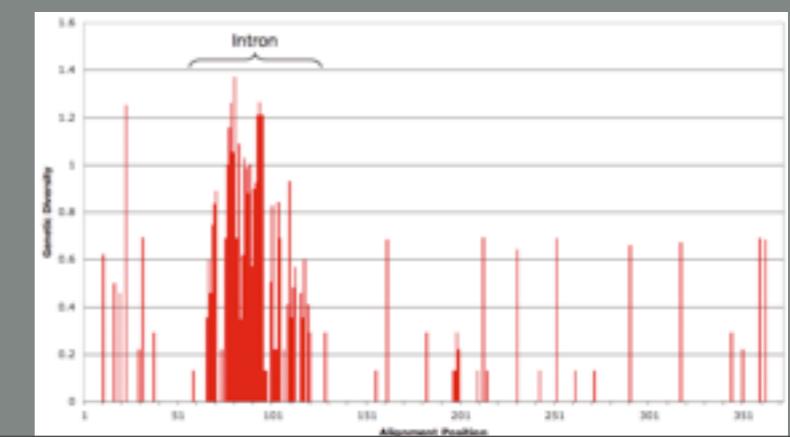
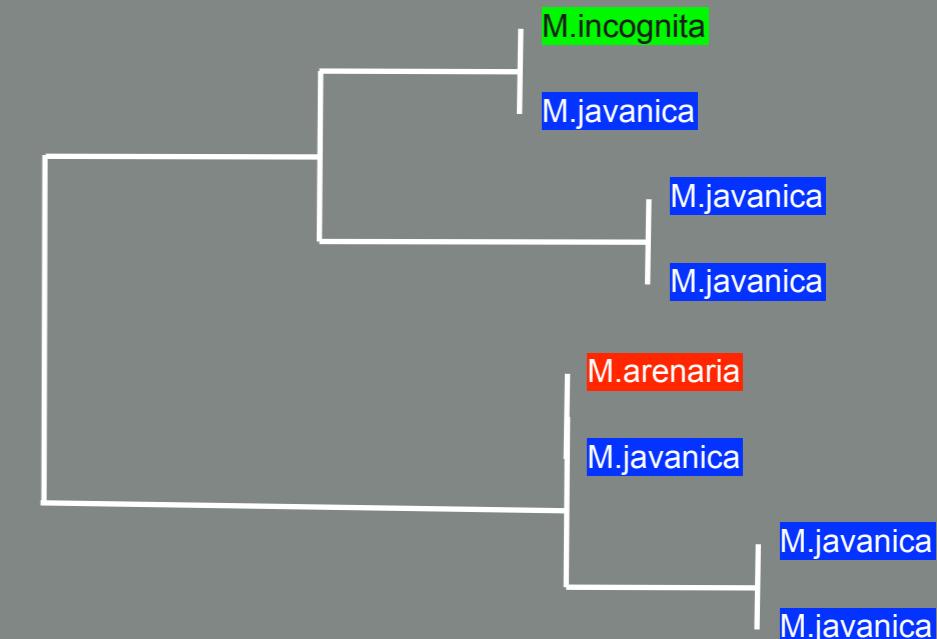


# Extreme allelic sequence divergence

Allelic Sequence Divergence levels are much greater in asexuals *Meloidogyne* than sexual *Meloidogyne*



RNA polymerase II		
Species	Max intraspecific substitutions	Substitutions to closest relative
<i>M. incognita</i>	15	0 <i>M. javanica</i>
<i>M. javanica</i>	16	0 <i>M. incognita</i>
Dystrophin		
Species	Max intraspecific substitutions	Substitutions to closest relative
<i>M. javanica</i>	30	0 <i>M. arenaria</i>
<i>M. arenaria</i>	32	0 <i>M. javanica</i>

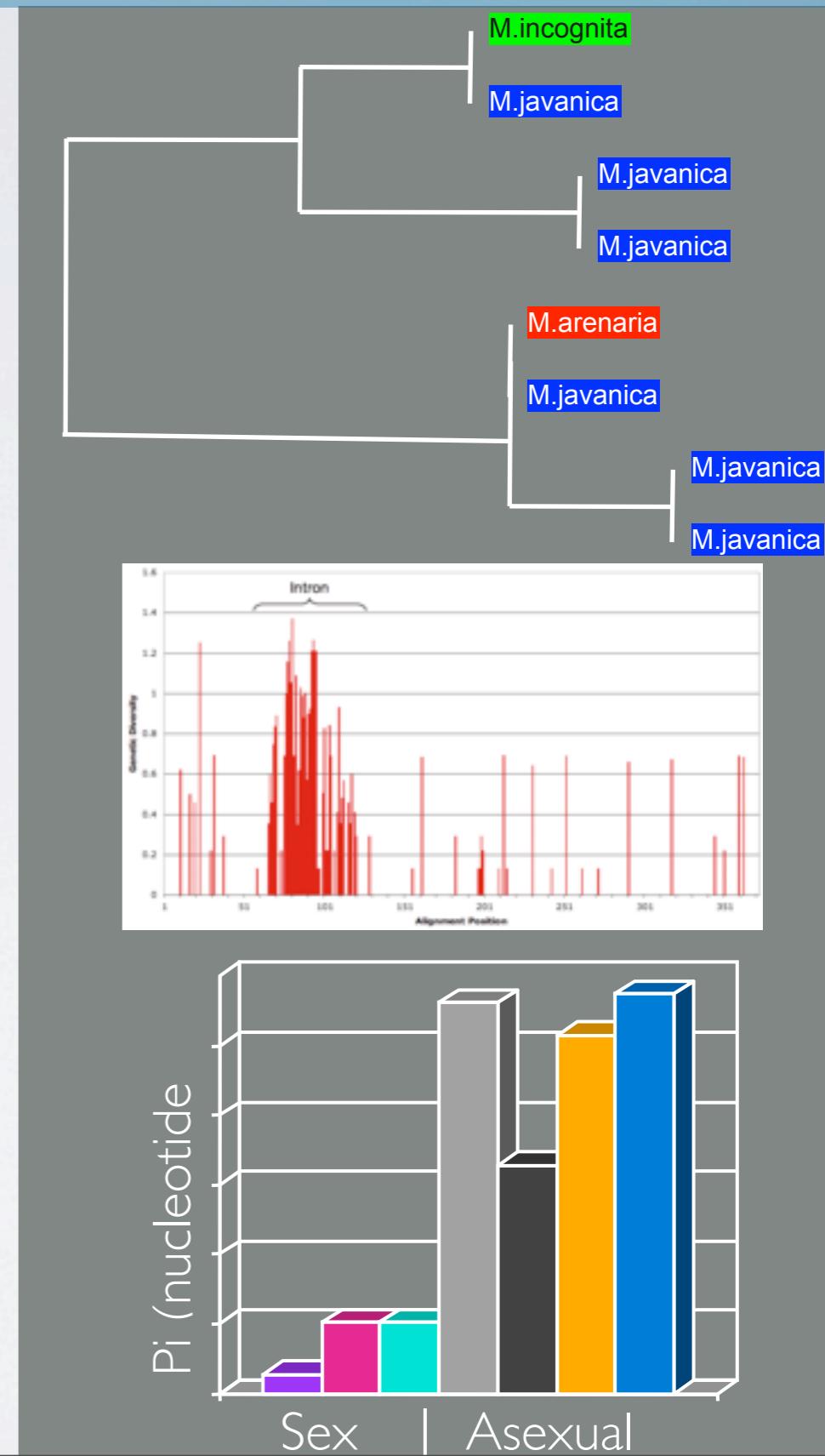


# Extreme allelic sequence divergence

ASD can be very large within asexual individuals

RNA polymerase II		
Species	Max intraspecific substitutions	Substitutions to closest relative
<i>M. incognita</i>	15	0 <i>M. javanica</i>
<i>M. javanica</i>	16	0 <i>M. incognita</i>
Dystrophin		
Species	Max intraspecific substitutions	Substitutions to closest relative
<i>M. javanica</i>	30	0 <i>M. arenaria</i>
<i>M. arenaria</i>	32	0 <i>M. javanica</i>

Yet identical alleles can be found between different species



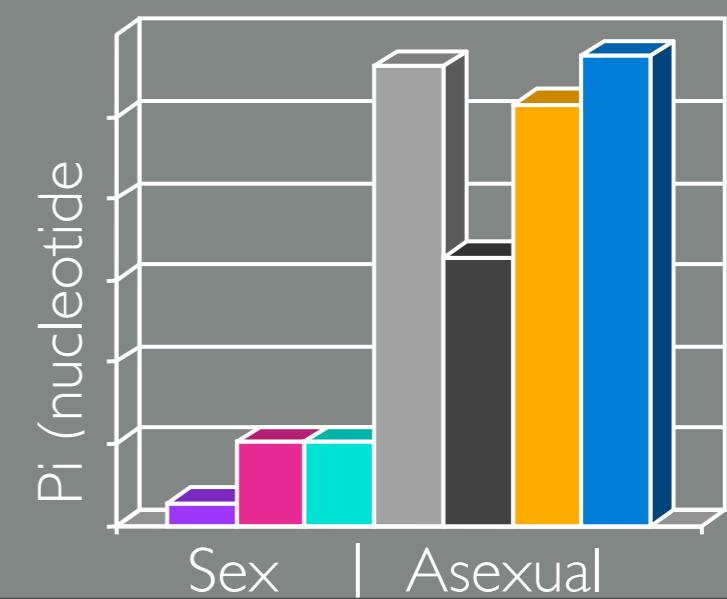
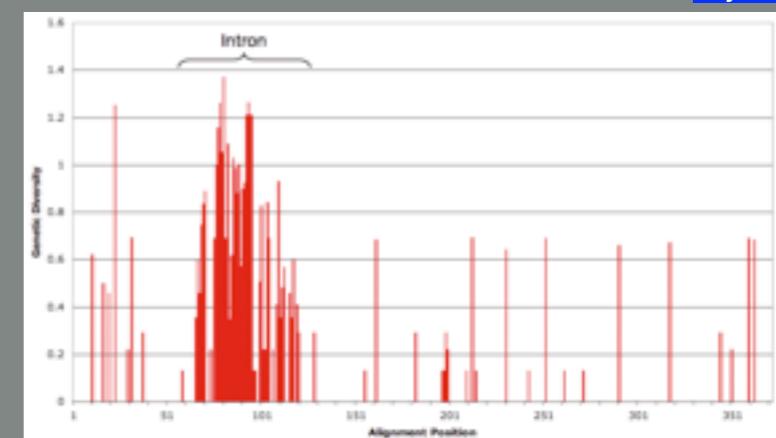
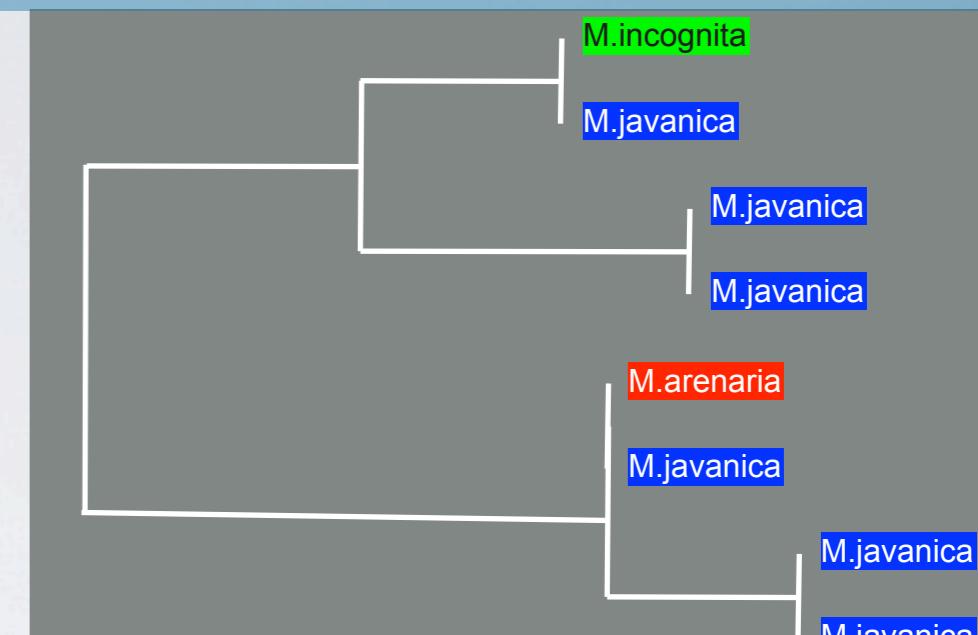
# Extreme allelic sequence divergence

RNA polymerase II

Species	Max intraspecific substitutions	Substitutions to closest relative
<i>M. incognita</i>	15	0 <i>M. javanica</i>
<i>M. javanica</i>	16	0 <i>M. incognita</i>

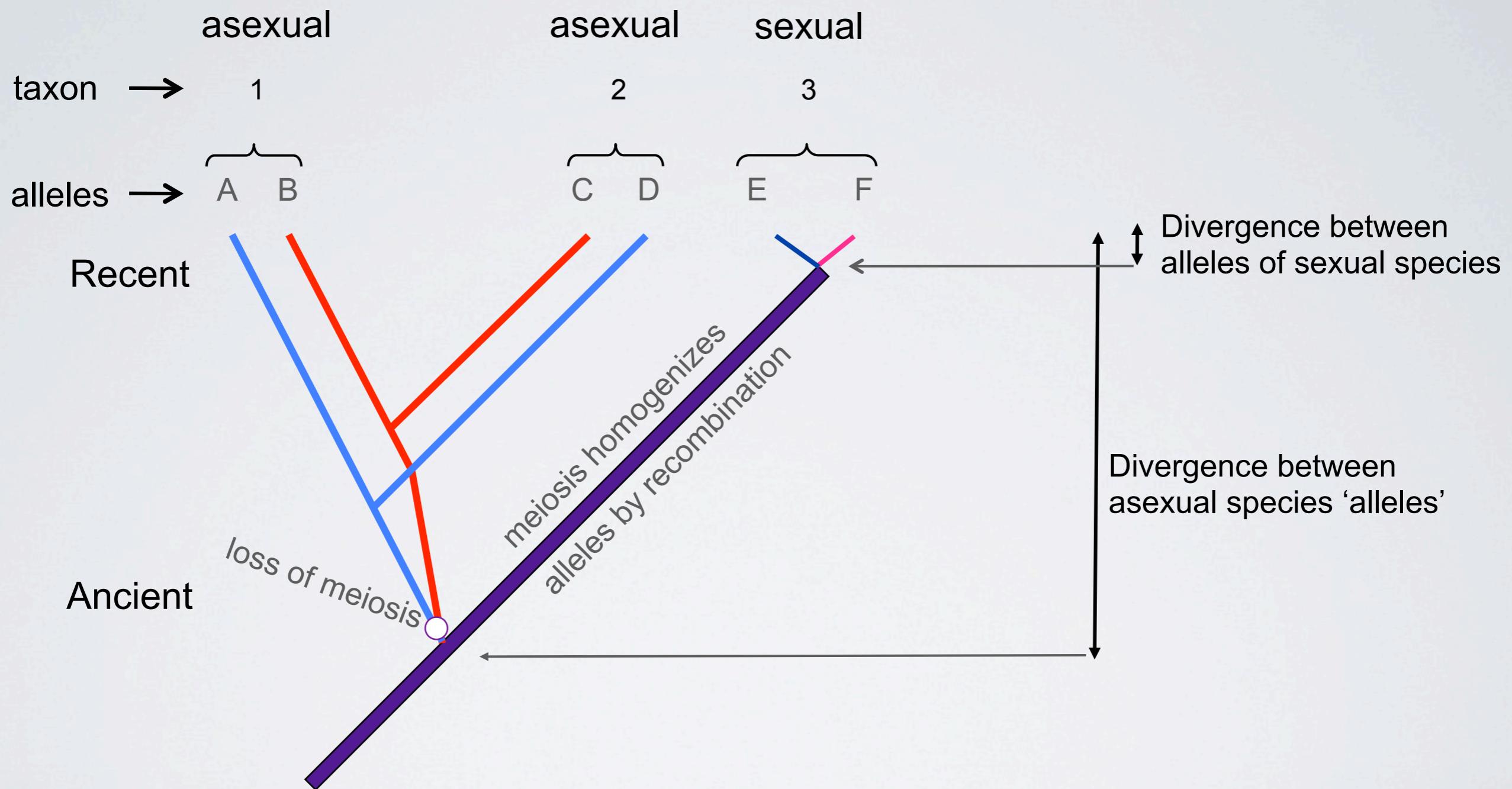
Dystrophin

Species	Max intraspecific substitutions	Substitutions to closest relative
<i>M. javanica</i>	30	0 <i>M. arenaria</i>
<i>M. arenaria</i>	32	0 <i>M. javanica</i>

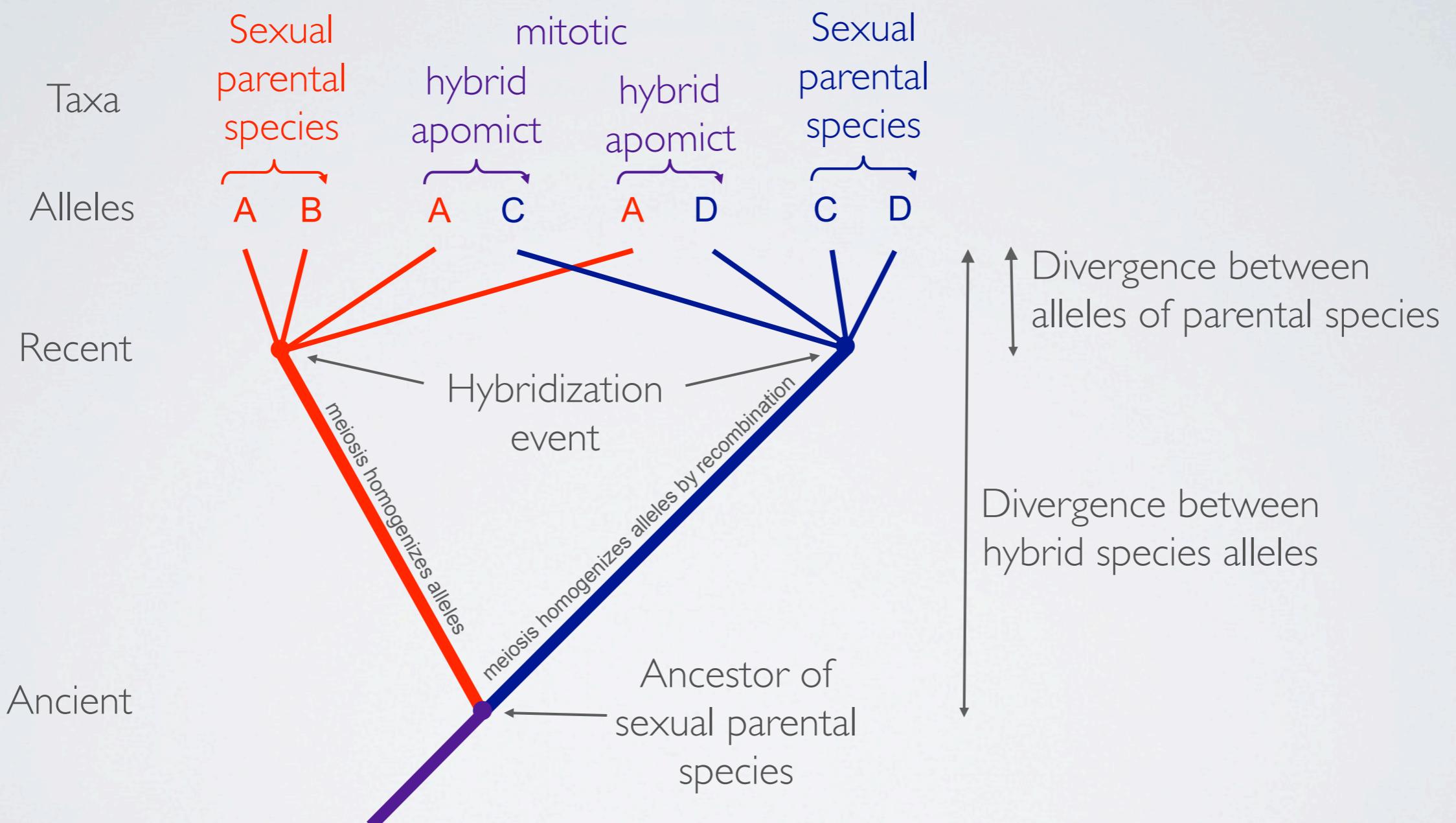


This allele sharing is not predicted by ancient asexuality, and suggests interspecific hybridization

# Extreme Asexual ASD



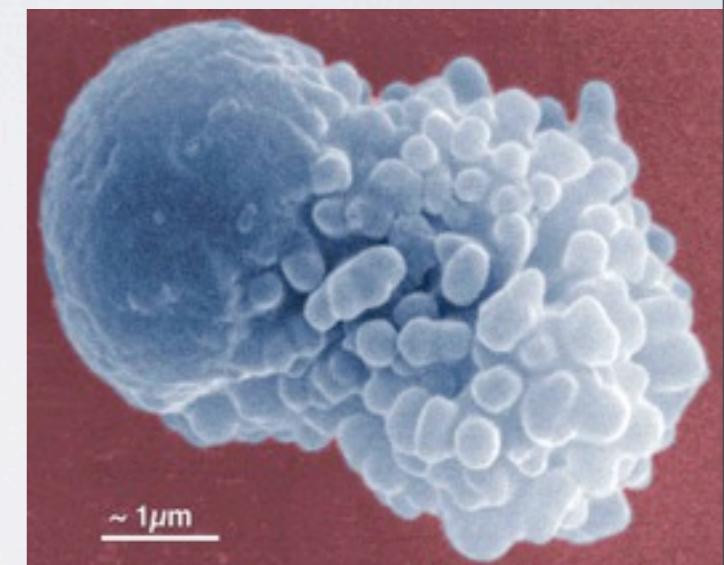
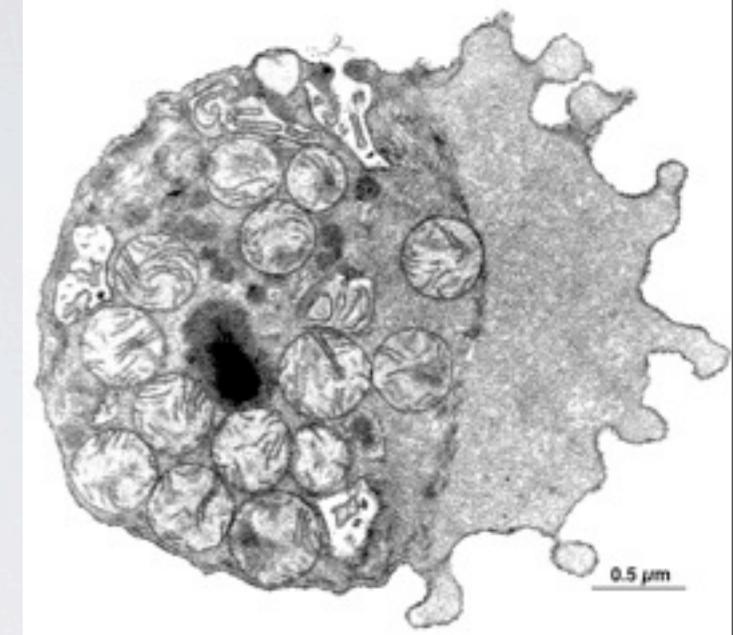
# Extreme Hybrid ASD



# Sex Specific Loci

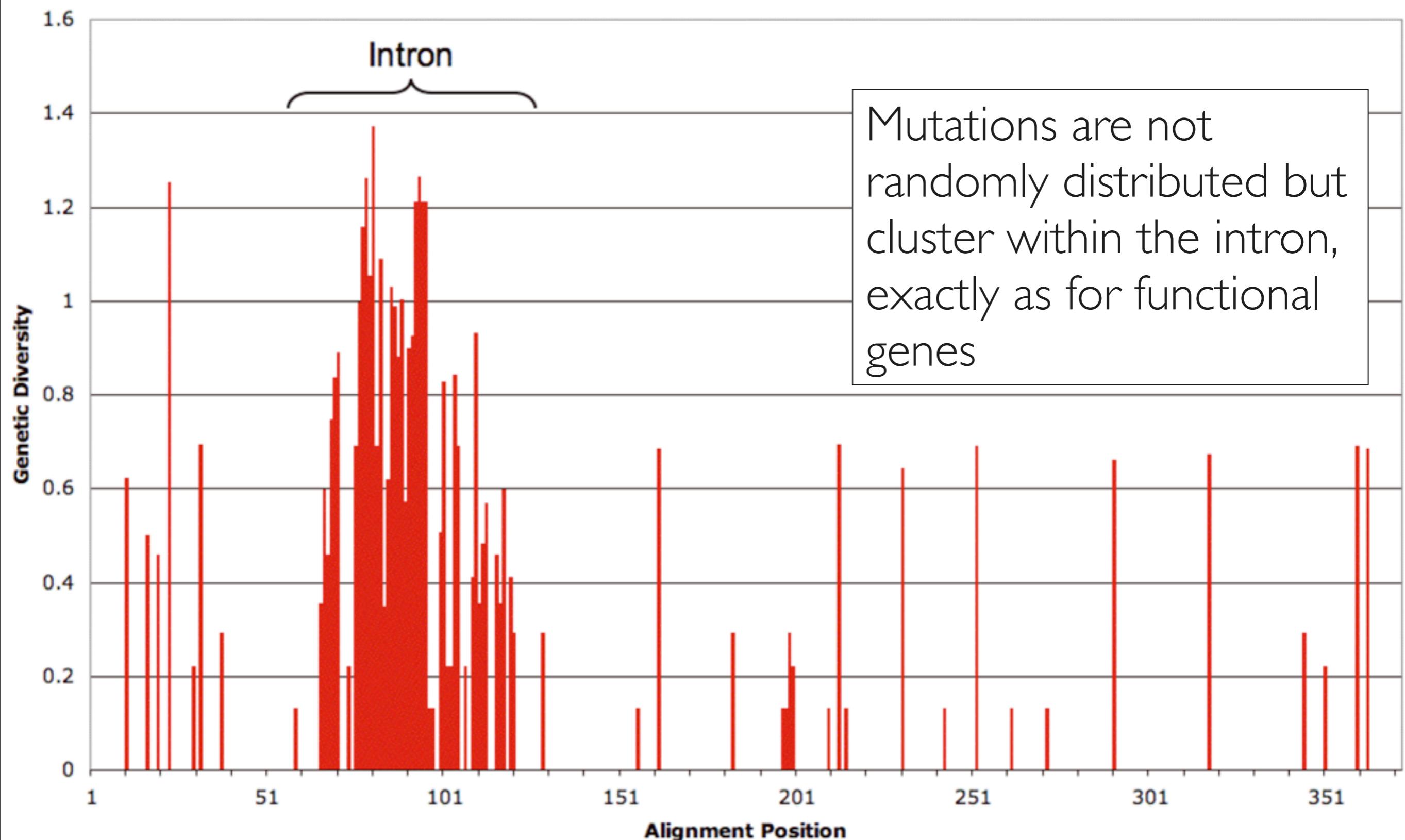
## Major Sperm Protein

- Nematodes have amoeboid (crawling) sperm
- *msp* genes only expressed in sperm and spermatocytes
  - Structural protein of sperm
  - Signal to recommence meiosis
- Prediction: *msp* gene should show signatures of loss of function (pseudogenization) in asexuals



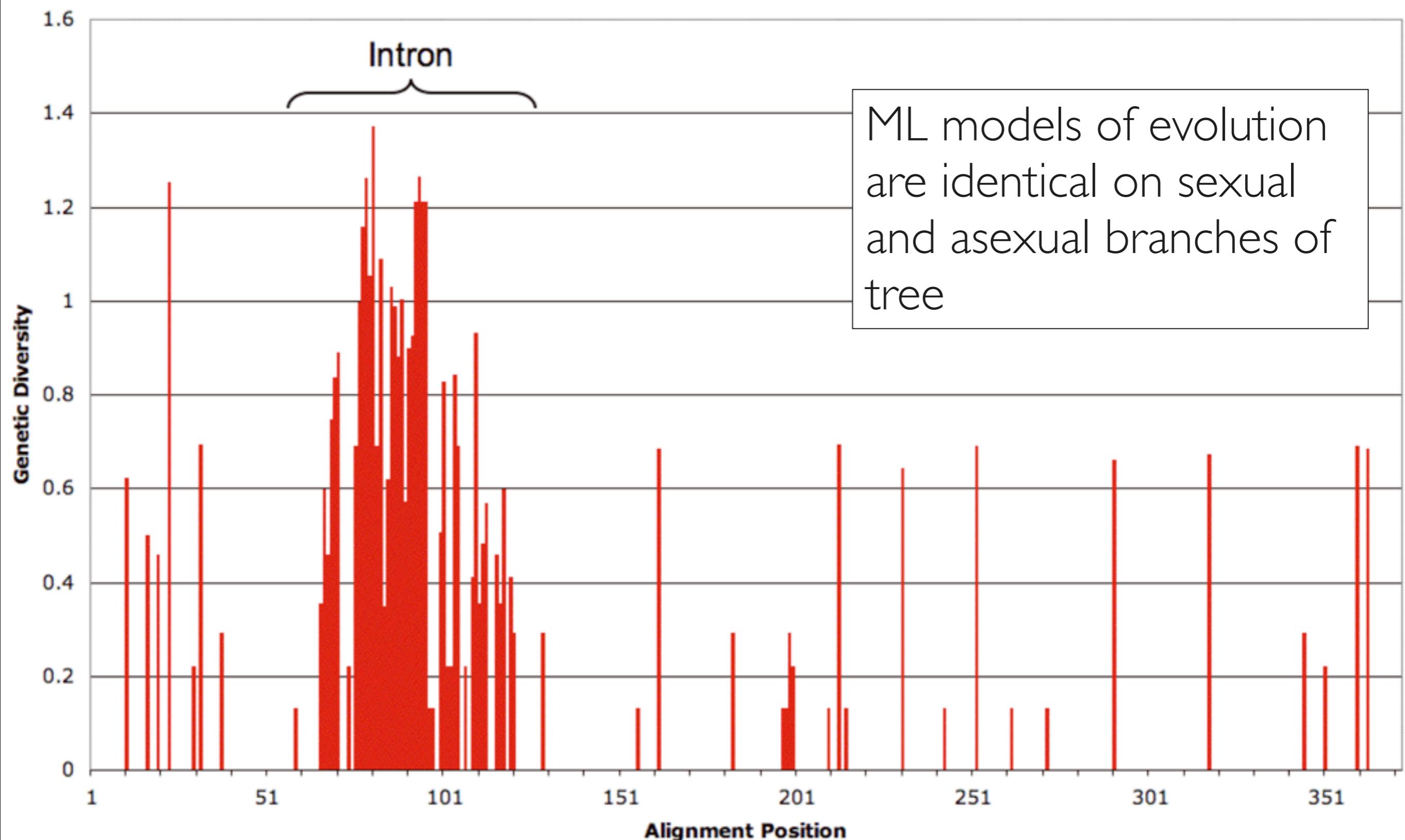
Electron microscope images from  
Ward lab, [http://  
www.mcb.arizona.edu/wardlab/](http://www.mcb.arizona.edu/wardlab/)

# msp intron diversity in asexuals



**Selection on this gene cannot have been abandoned anciently**

# msp intron diversity in asexuals



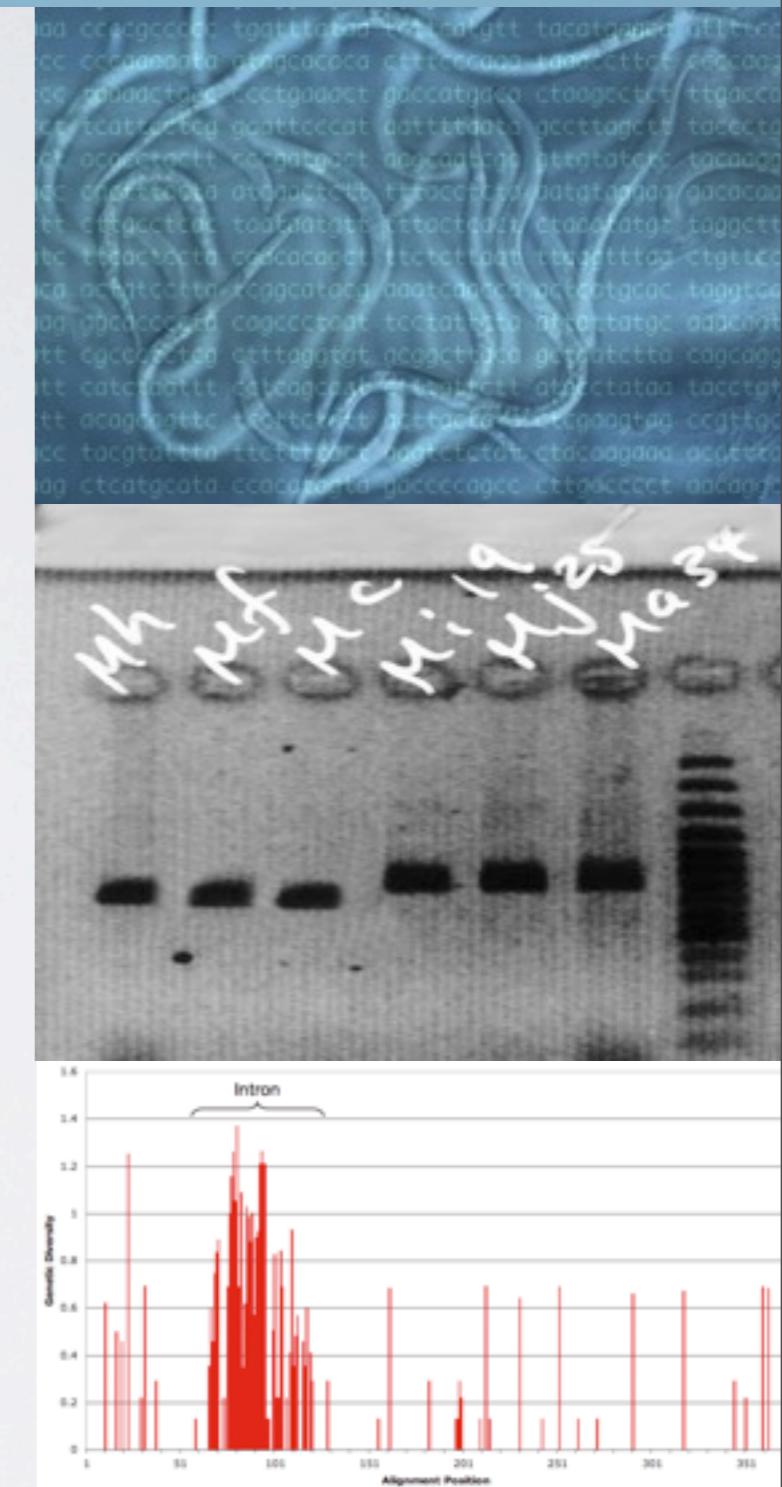
**Selection on this gene cannot have been abandoned anciently**

# Previous Single Gene Sequencing

Lunt DH 2008 BMC Evolutionary Biology 8:194

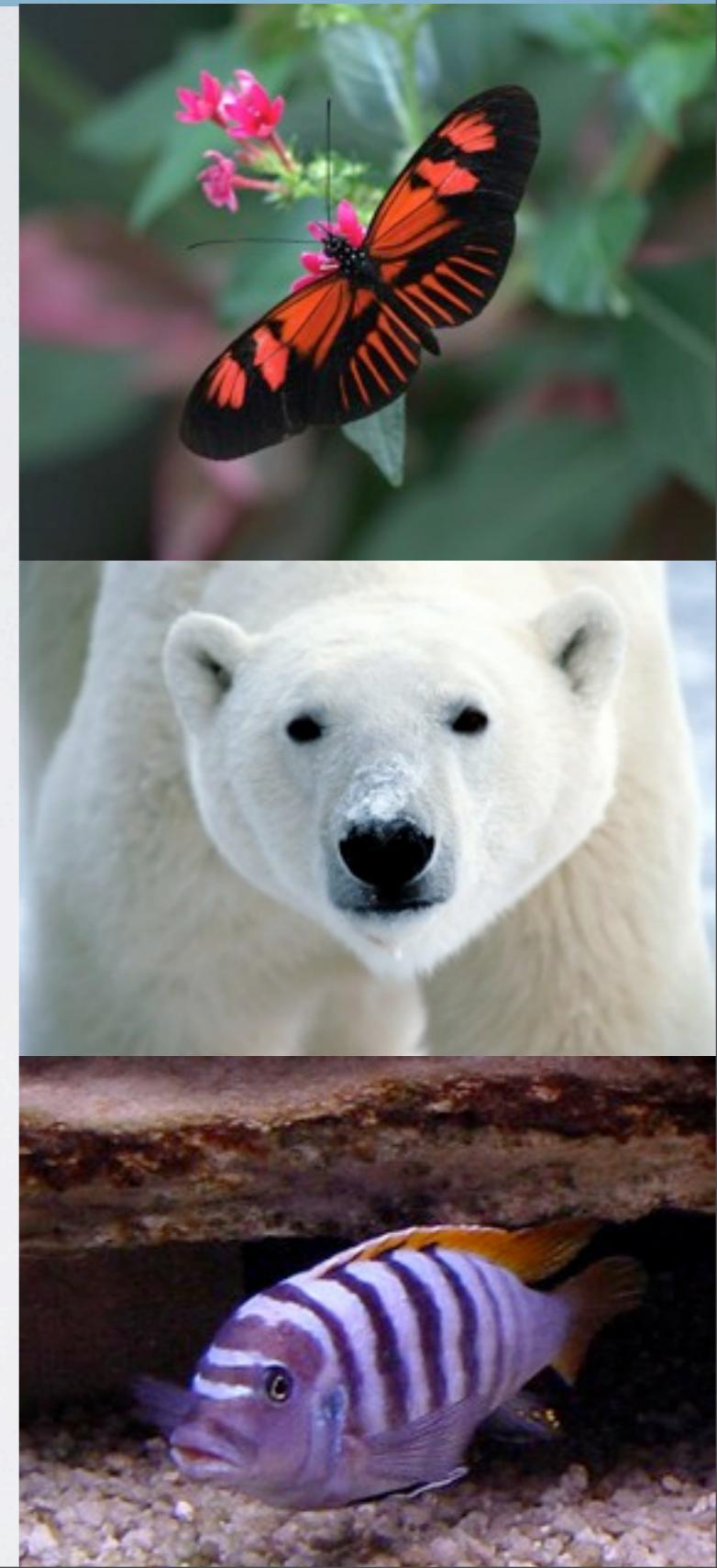
I can reject *ancient asexuality* on basis of interspecific allele sharing and identical molecular evolution of sperm protein genes

Data suggests interspecific hybrid origins



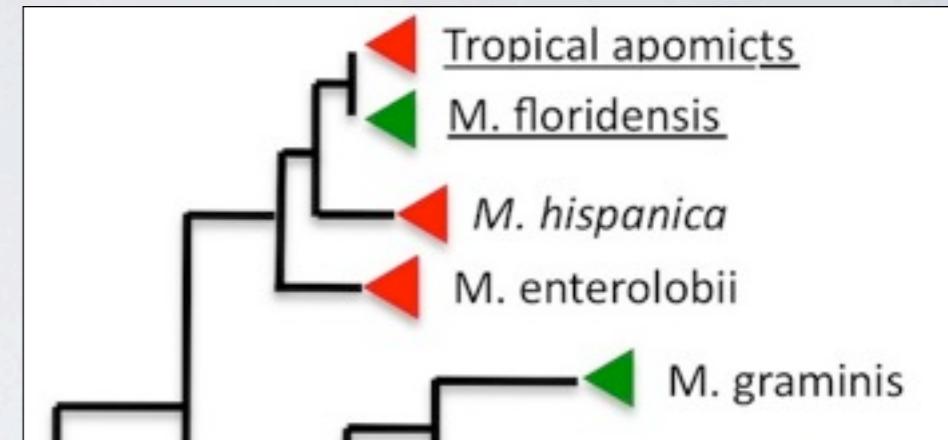
# Hybrid Speciation

- Once thought that hybrid speciation was rare and inconsequential in animals
- Genome biology is revealing a very different view
- We have investigated the origins of *Meloidogyne* asexuals in this context



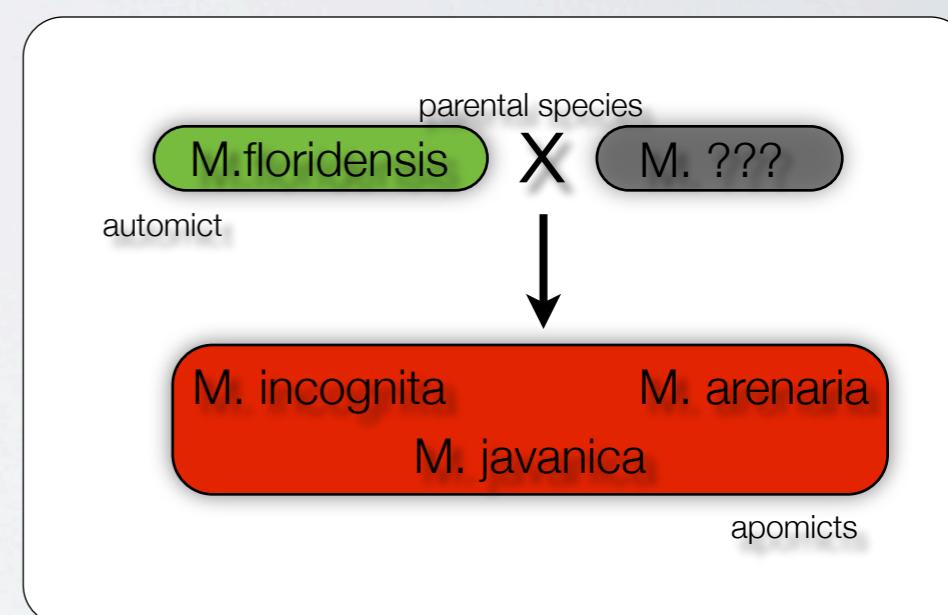
# Is *M. floridensis* the parent of the asexuals?

*M. floridensis* is found within the phylogenetic diversity of asexual species



It reproduces sexually by automixis

Could it be a parent of the asexual lineages via interspecific hybridization?

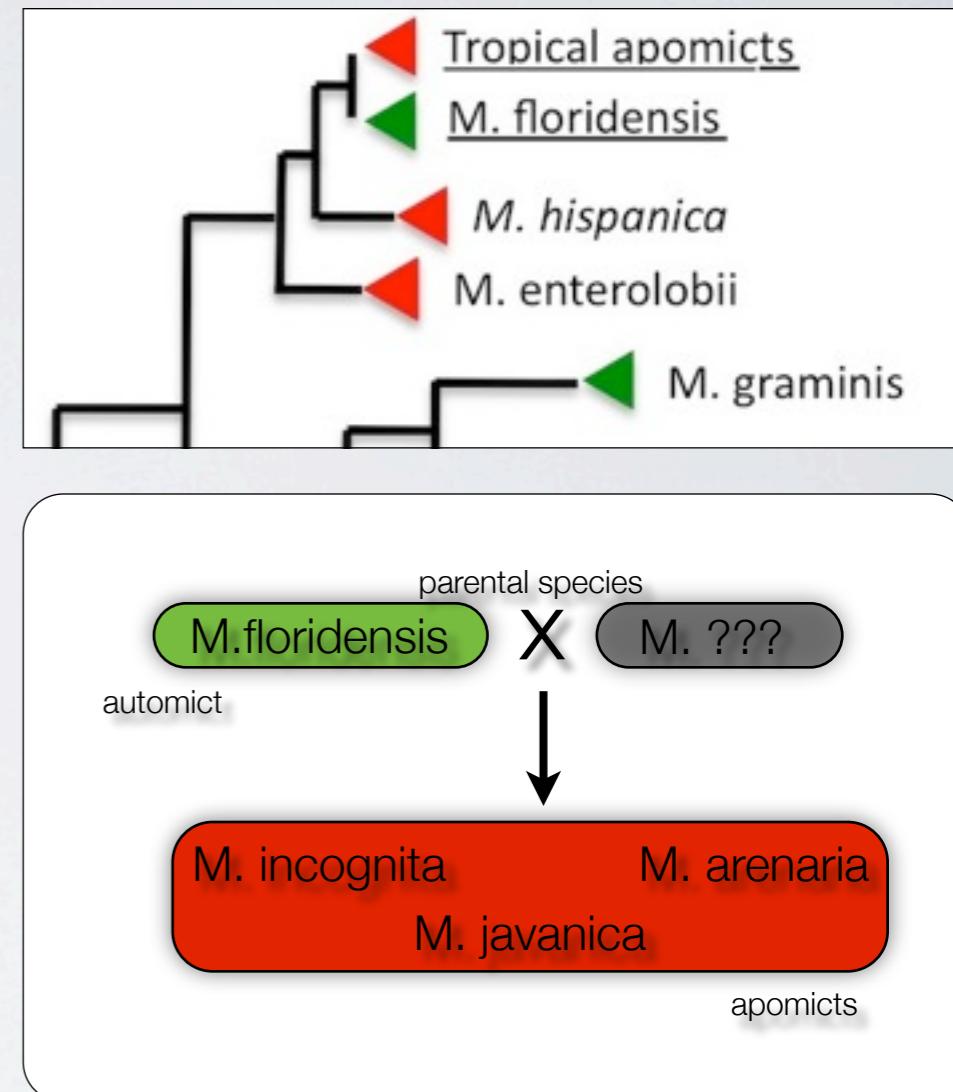


# Is *M. floridensis* the parent of the asexuals?

Investigated using whole genome sequences and 2 distinct approaches;

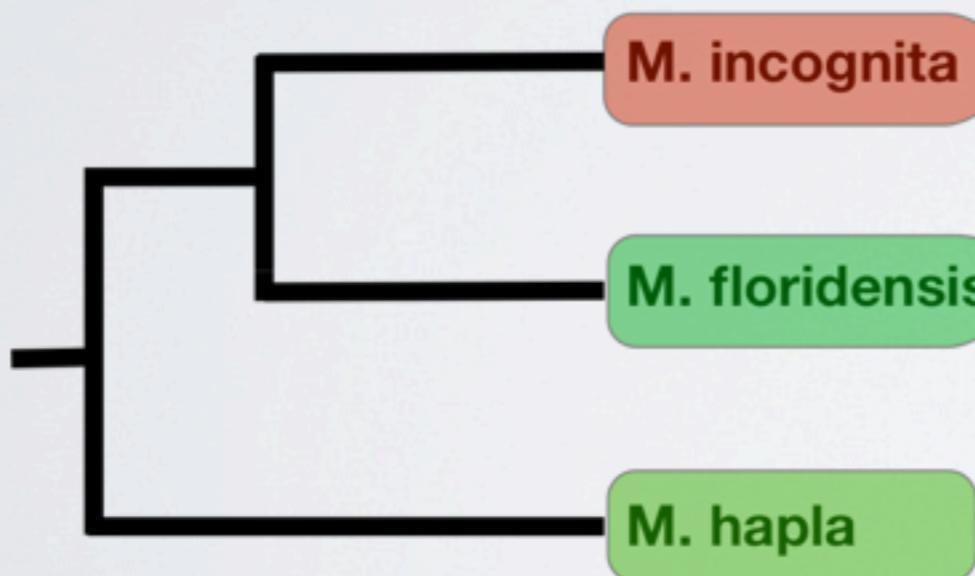
--look at the within-genome patterns of diversity to determine hybrid nature of genomes

--look at phylogenetic relationships of all genes to study origins and parents



# Meloidogyne comparative genomics

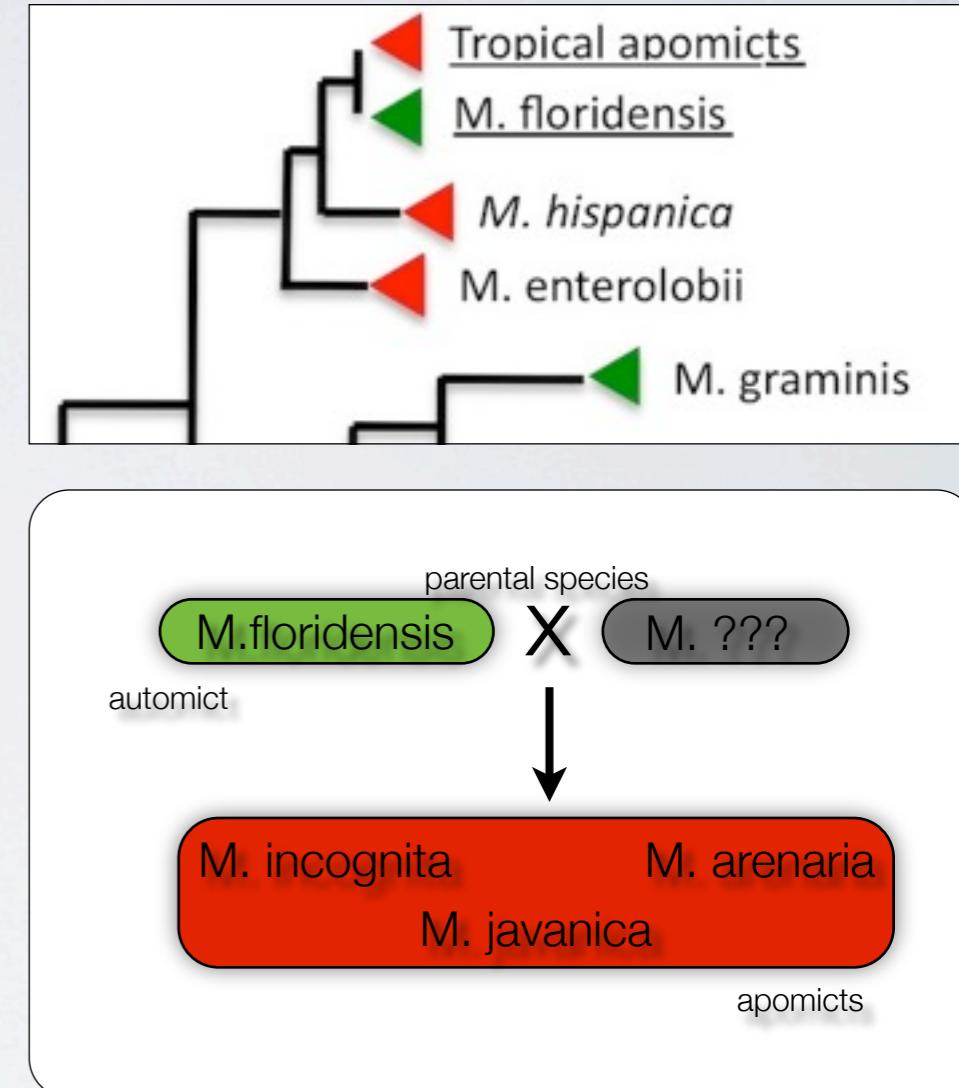
We have sequenced *M. floridensis* genome and are able to compare to 2 other *Meloidogyne* genomes published by other groups



asexual, hybrid?

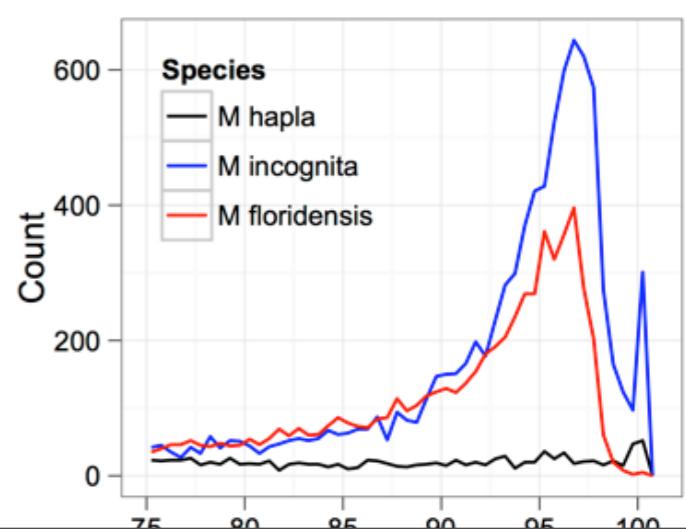
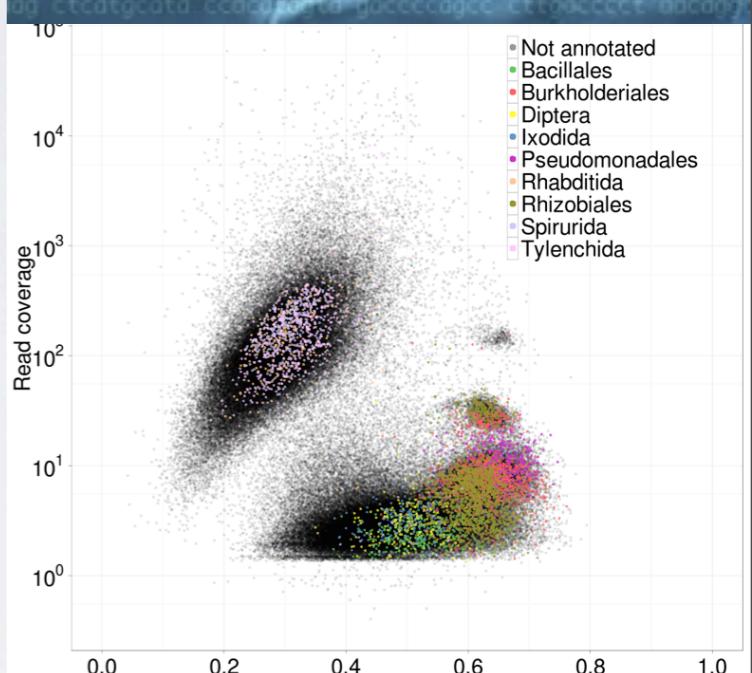
sexual, parental?

sexual, outgroup



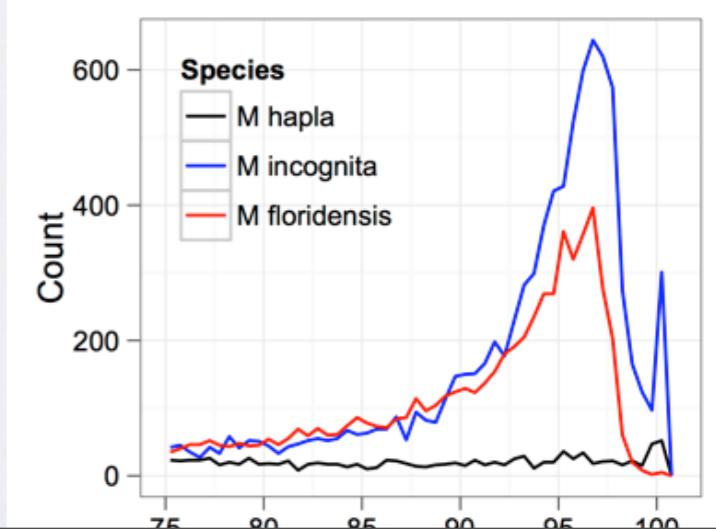
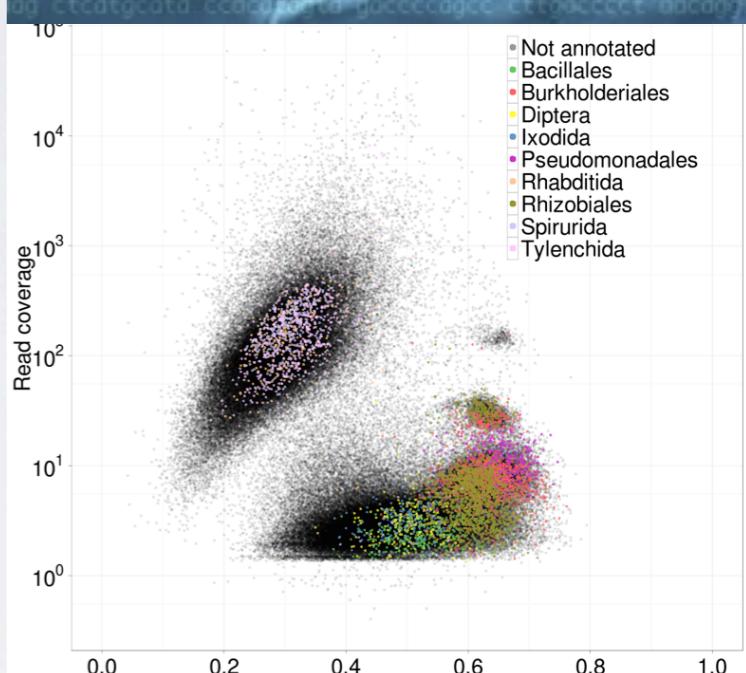
# The *Meloidogyne floridensis* genome

- 100Mb assembly ~100x genomic coverage
- 15.3k predicted proteins
- Directly comparable to published *Meloidogyne* genomes



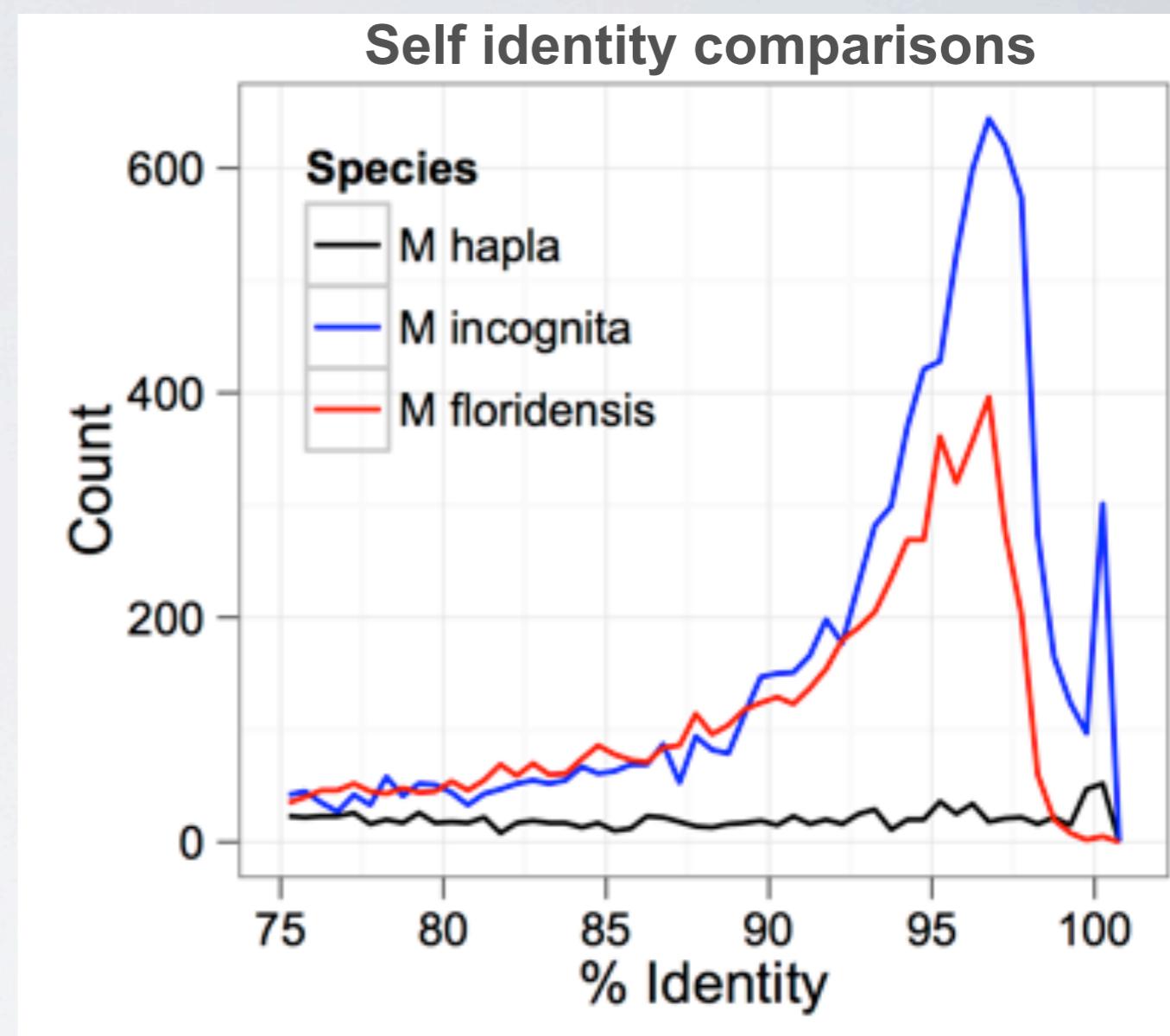
# Comparative genomics questions

- Is there evidence of hybrid origins of asexual species?
- Nature of hybridization?
- Is *M. floridensis* a parental?
- How do offspring and parental genomes differ?



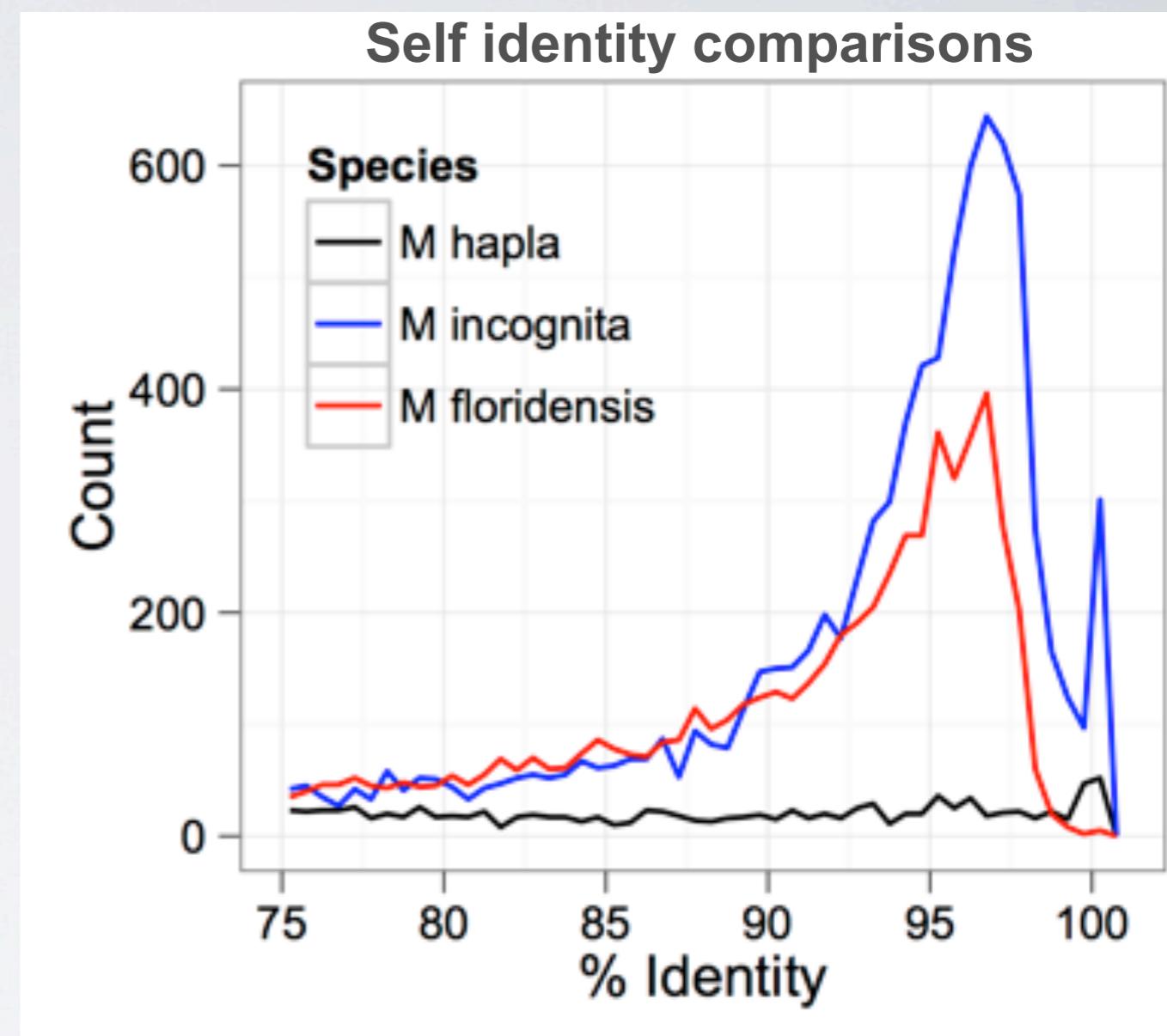
# ID of duplicated protein-coding regions

- Coding sequences from each of the three target genomes (*M. hapla*, *M. incognita* and *M. floridensis*) were compared to the set of genes from the same species
- The percent identity of the best matching (non-self) coding sequence was calculated, and is plotted as a frequency histogram
- Both *M. incognita* and *M. floridensis* show evidence of presence of many duplicates, while *M. hapla* does not



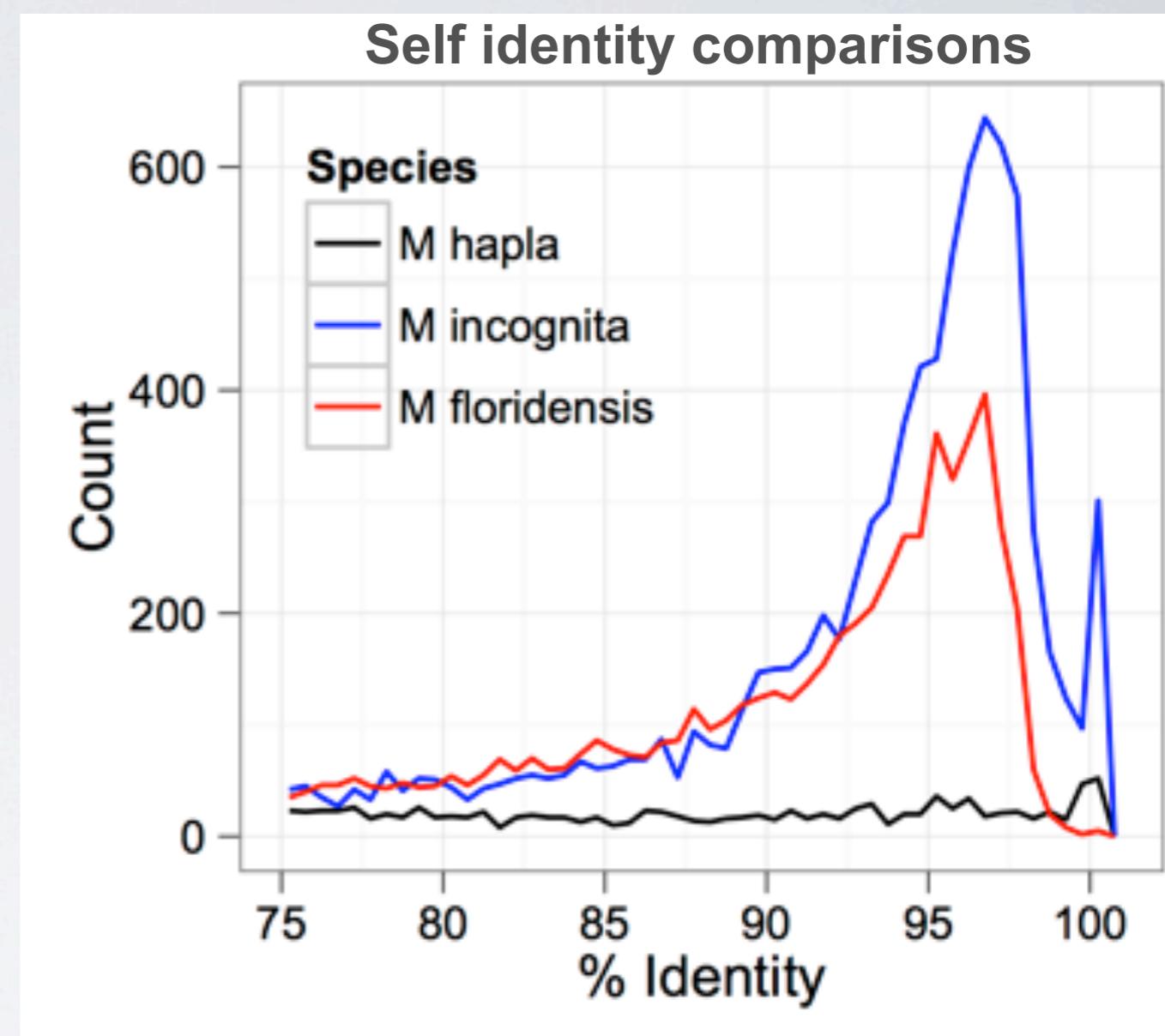
# ID of duplicated protein-coding regions

- Coding sequences from each of the three target genomes (*M. hapla*, *M. incognita* and *M. floridensis*) were compared to the set of genes from the same species
- The percent identity of the best matching (non-self) coding sequence was calculated, and is plotted as a frequency histogram
- Both *M. incognita* and *M. floridensis* show evidence of presence of many duplicates, while *M. hapla* does not



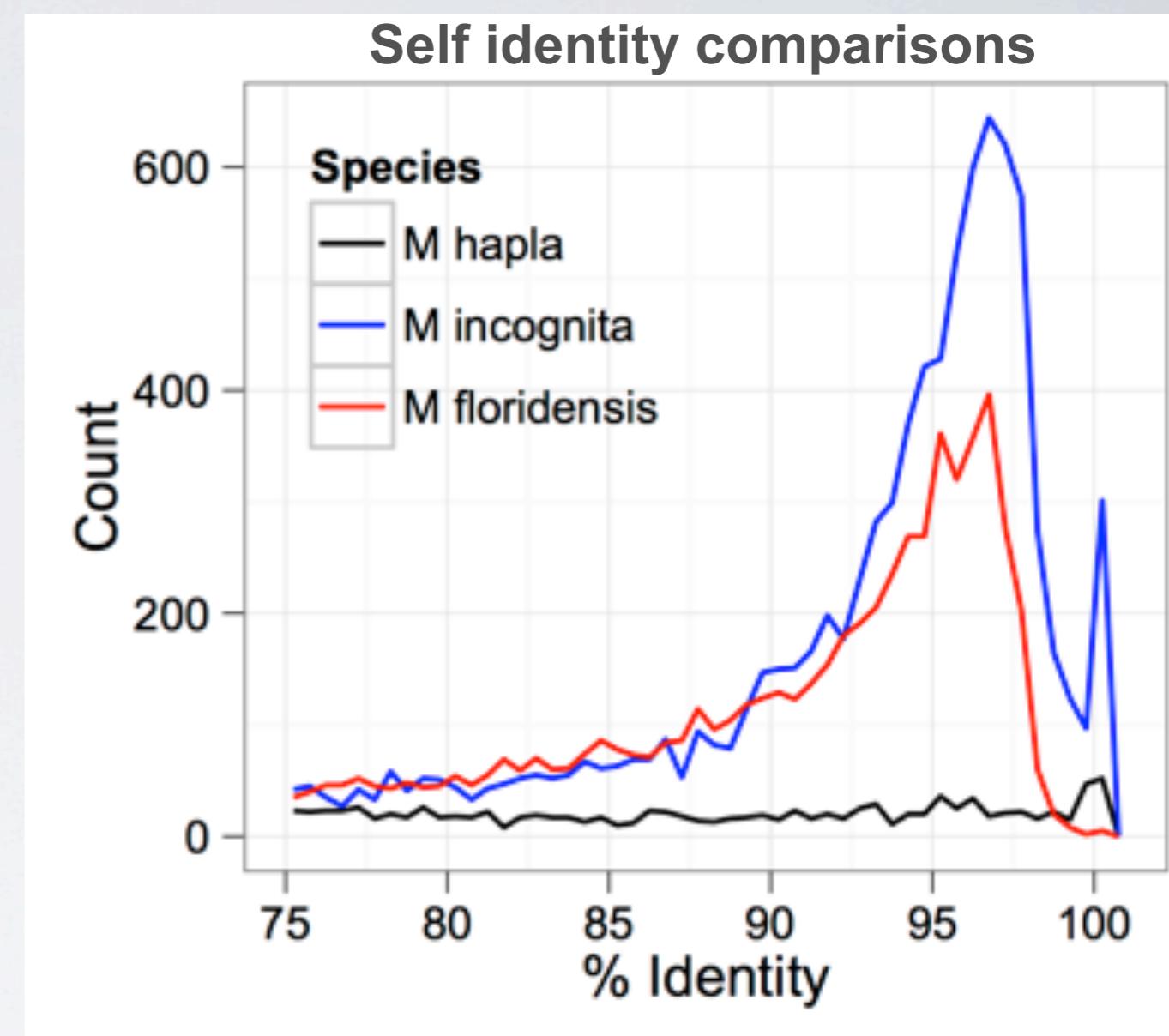
# ID of duplicated protein-coding regions

- Coding sequences from each of the three target genomes (*M. hapla*, *M. incognita* and *M. floridensis*) were compared to the set of genes from the same species
- The percent identity of the best matching (non-self) coding sequence was calculated, and is plotted as a frequency histogram
- Both *M. incognita* and *M. floridensis* show evidence of presence of many duplicates, while *M. hapla* does not



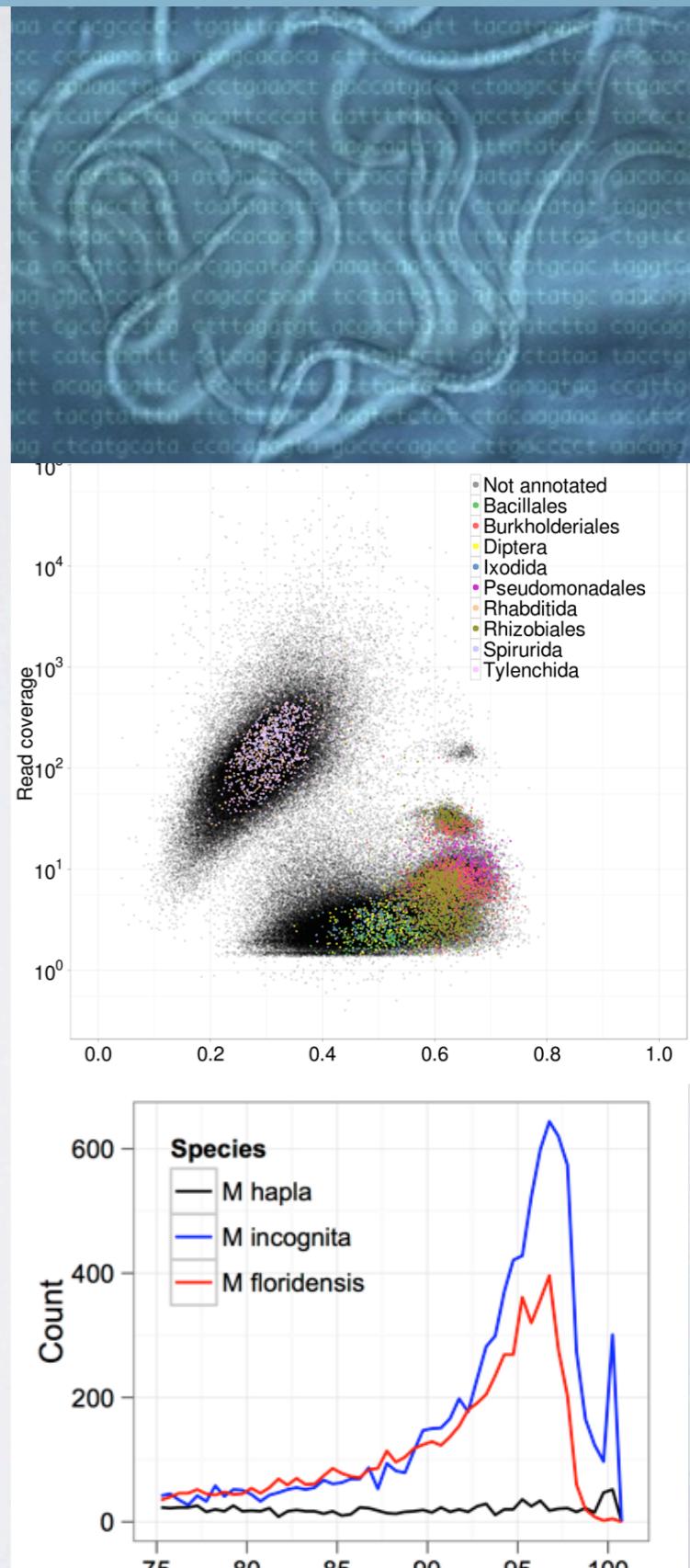
# ID of duplicated protein-coding regions

- Both *M. incognita* and *M. floridensis* contain diverged gene copies.
  - These loci duplicated at approximately the same point in time.
  - A ploidy change is not involved.
- This is expected pattern for hybrid genomes



# *M. floridensis* Genome Size

- Assembly size is not haploid genome size for hybrid species
- Divergence (4-8%) between homeologous (hybrid) copies will preclude assembly
- Our assembly of 100Mb is ~2x 50-54Mb genome size of *M. hapla*

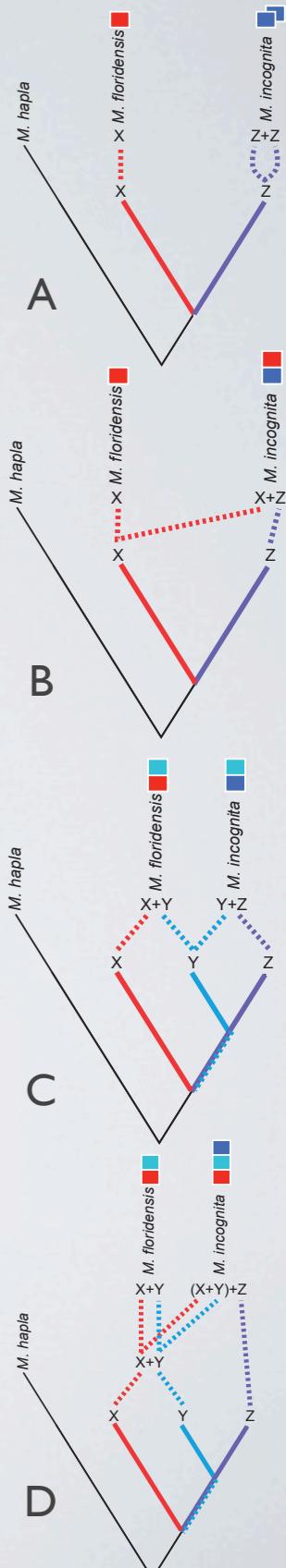


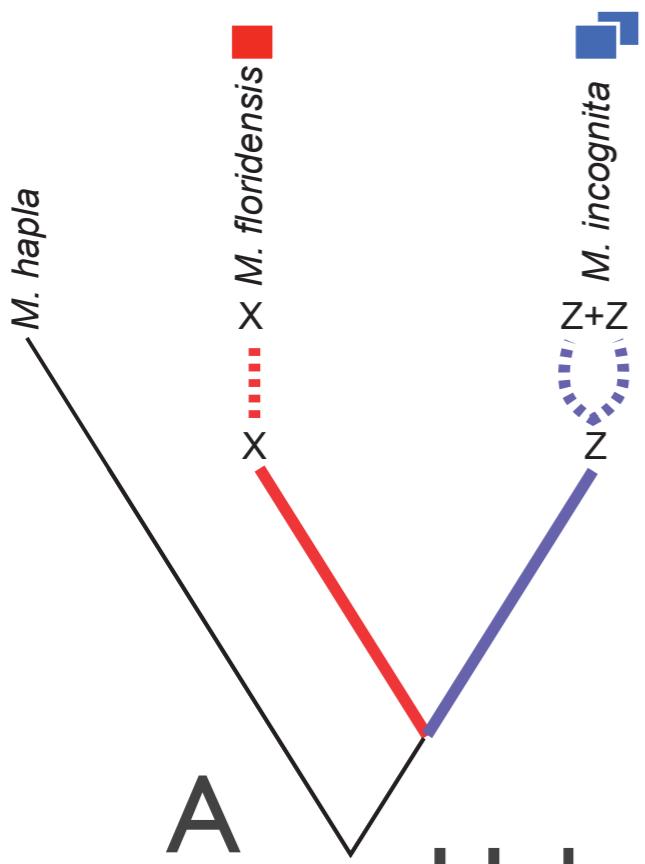
# Hybridization Hypotheses

There are very many ways species could hybridize, duplicate genes, lose genes

We have selected a broad range of possibilities informed by prior knowledge

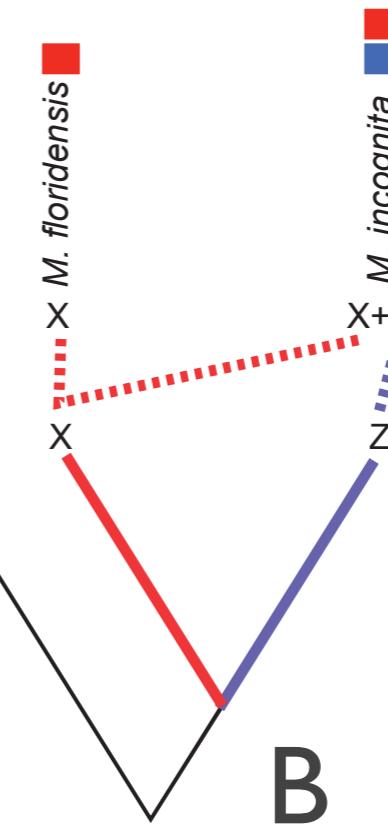
We have tested their predictions phylogenetically



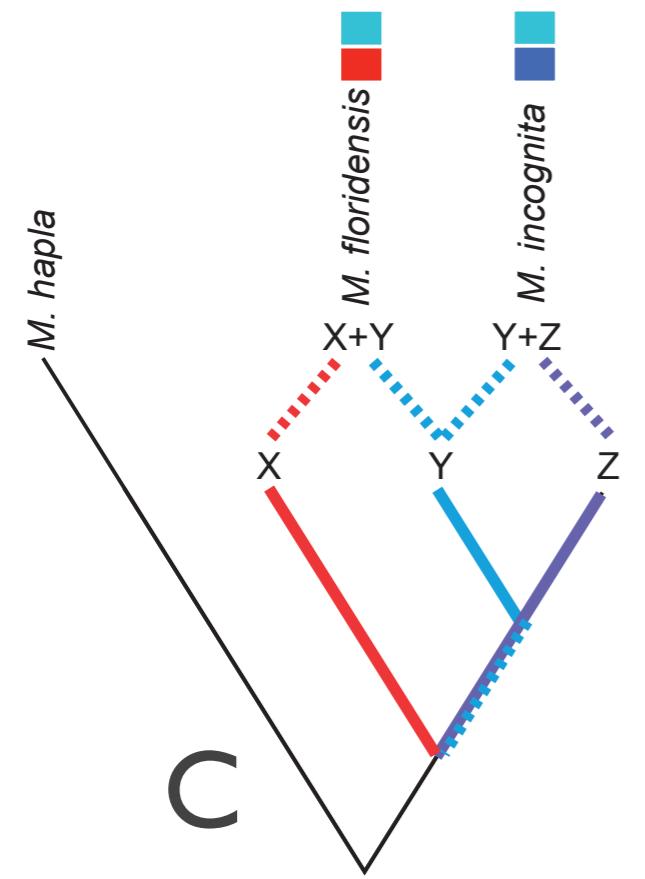


A

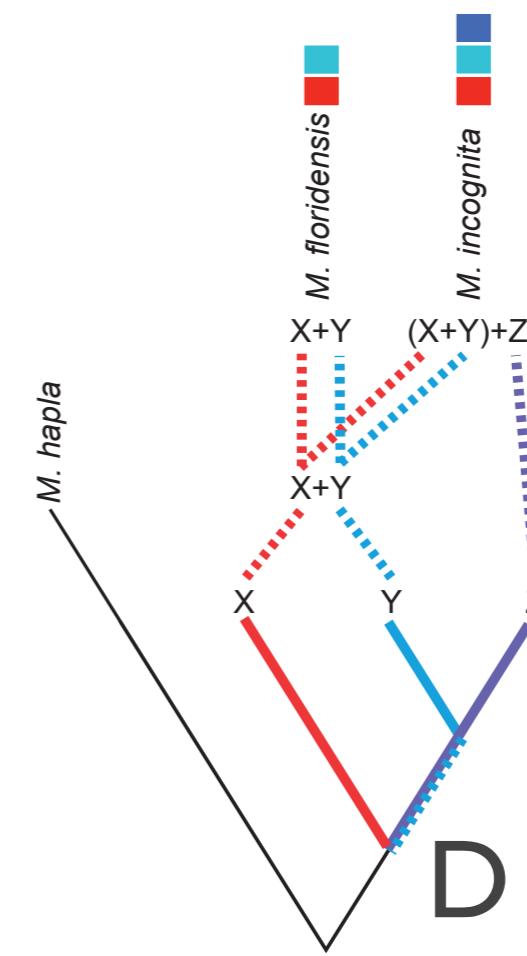
## Hybridization hypotheses



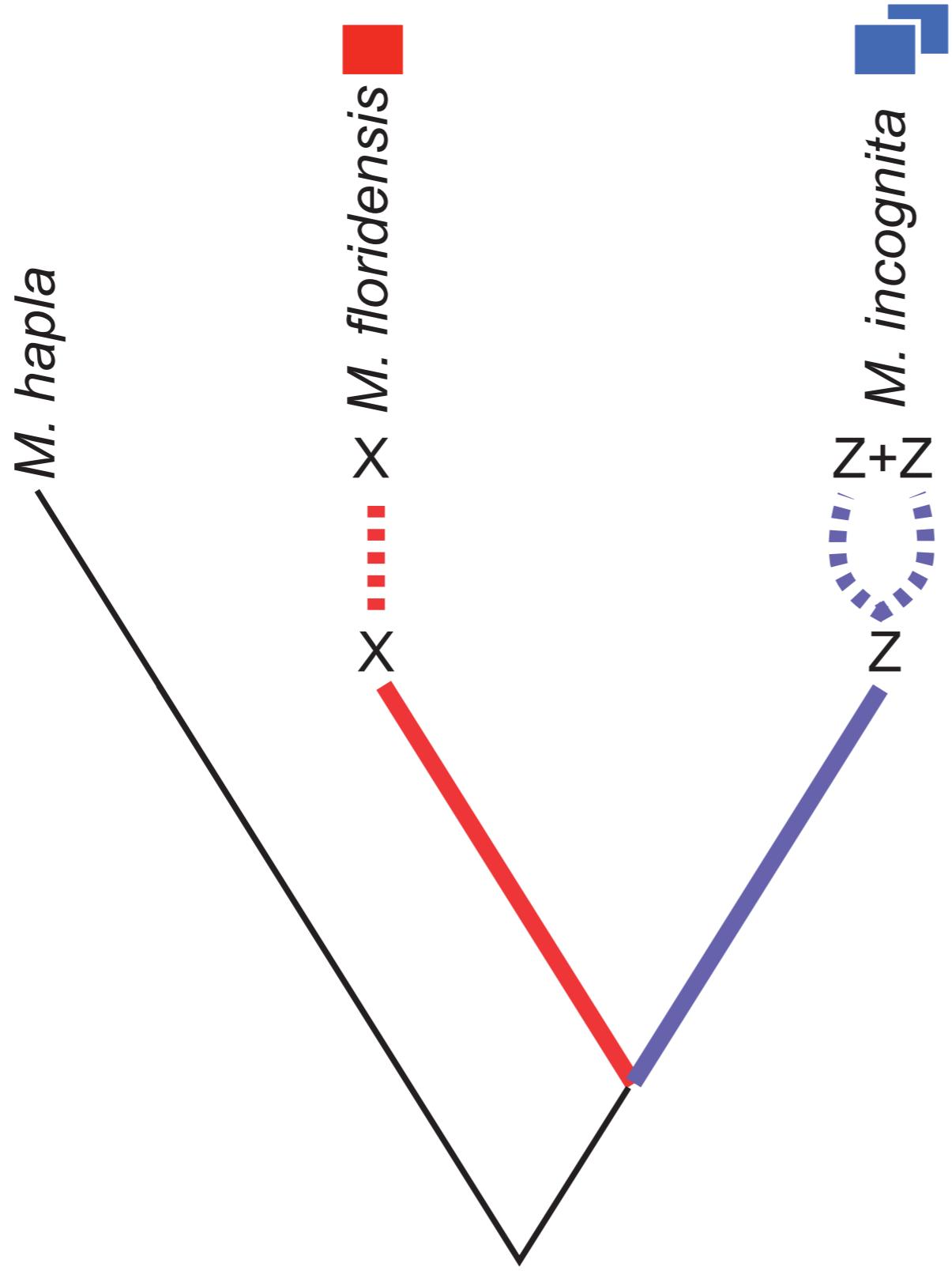
B



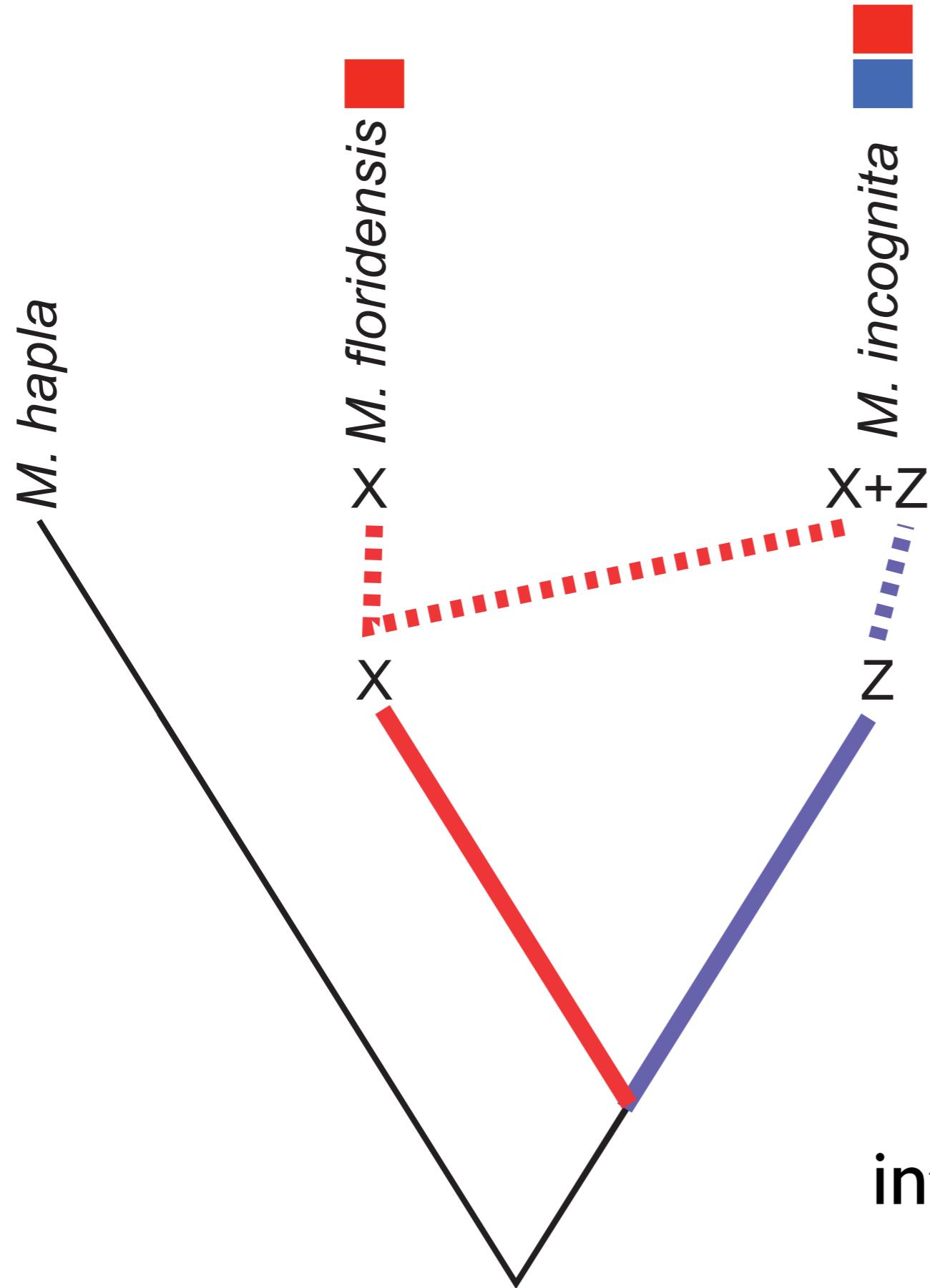
C



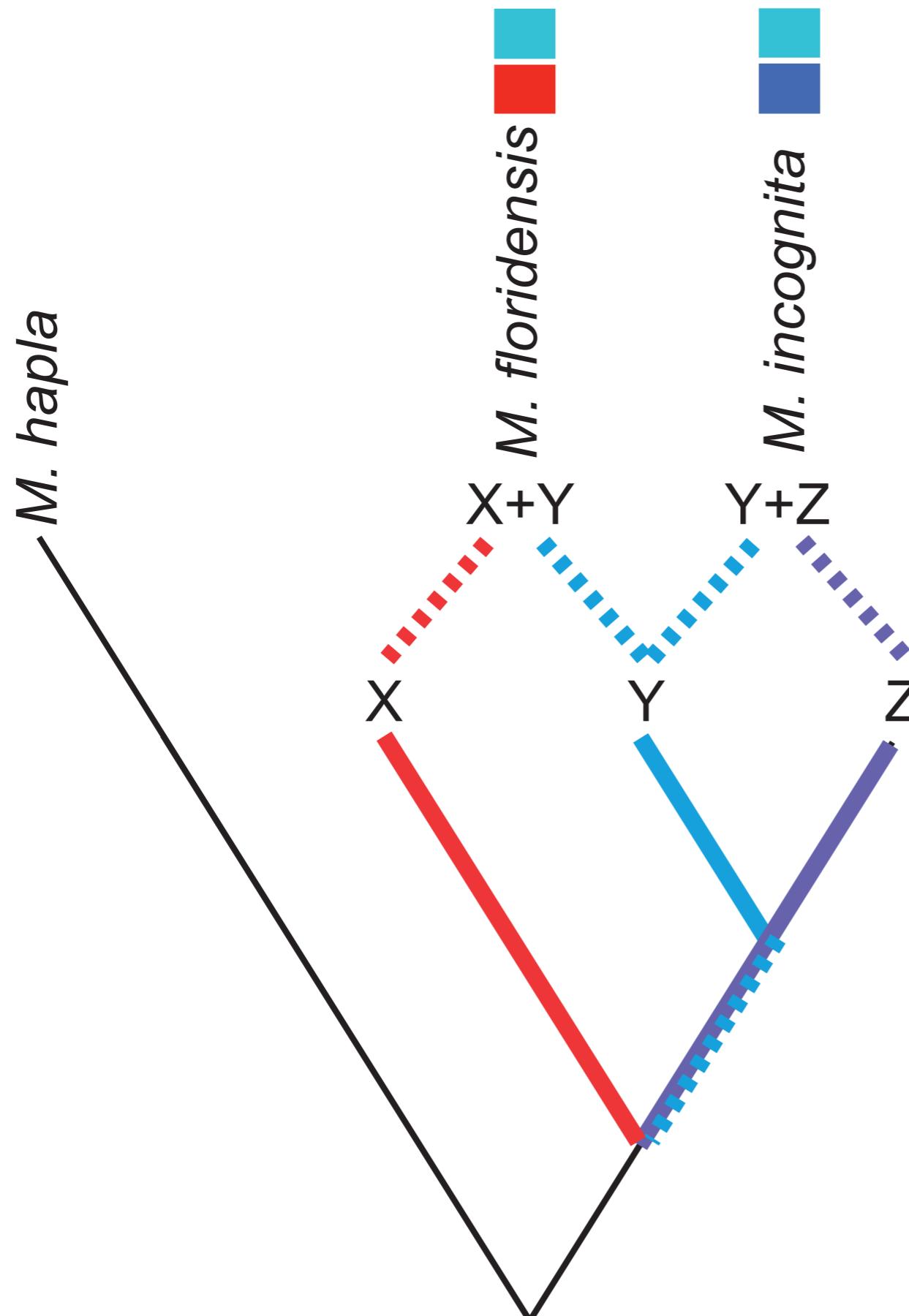
D



(A)  
Whole genome  
duplication(s)

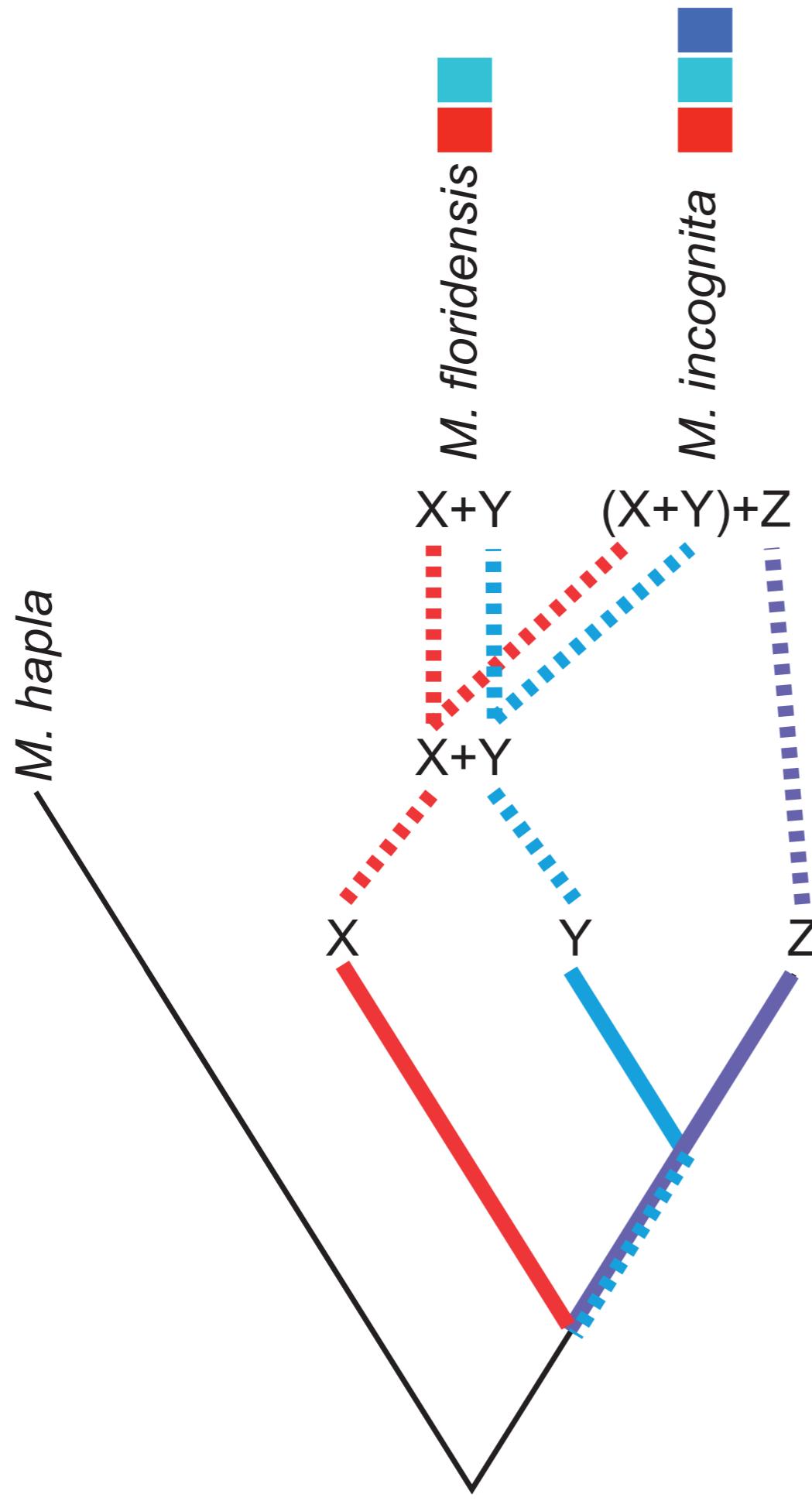


(B)  
*M. incognita* is an  
interspecific hybrid with  
*M. floridensis* as one  
parent



(C)

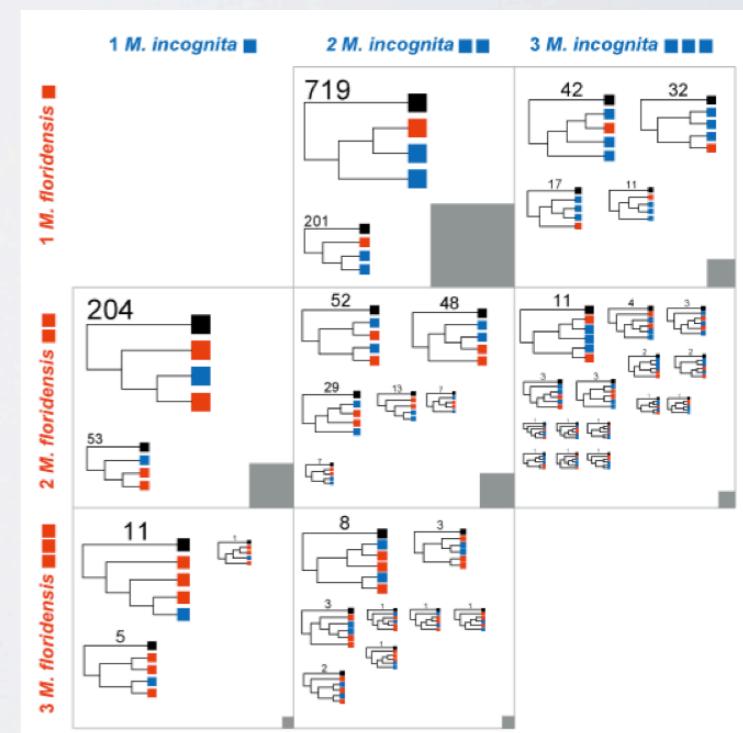
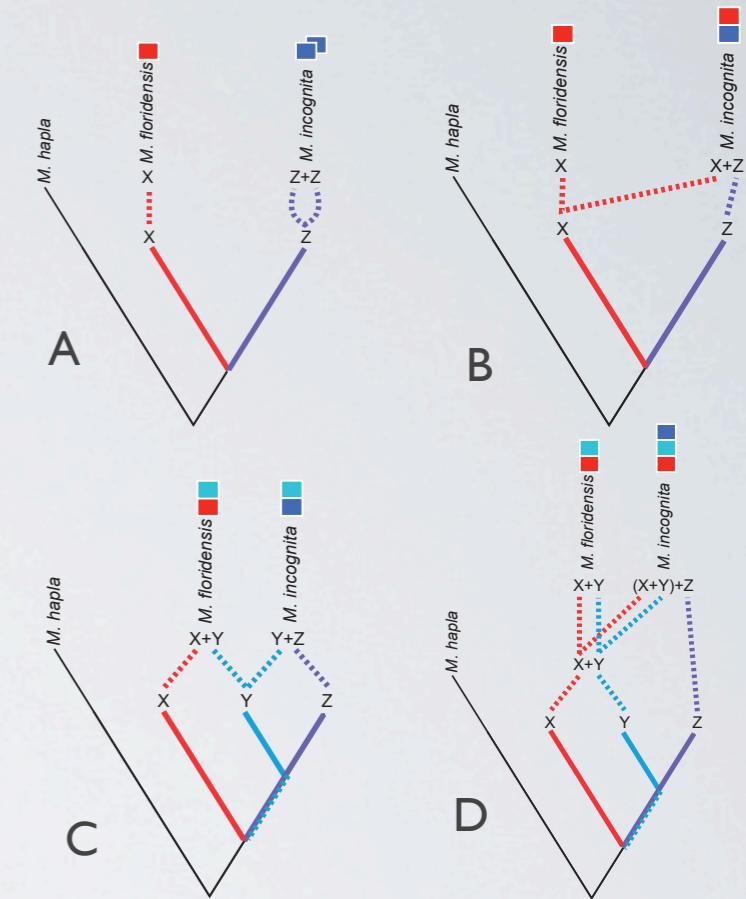
*M. incognita* and *M. floridensis* are independent hybrids sharing one parent



(D)  
*M. floridensis* is a hybrid  
 and *M. incognita* is a  
 secondary hybrid  
 between *M. floridensis* and  
 a 3<sup>rd</sup> parent

# Testing by Phylogenomics

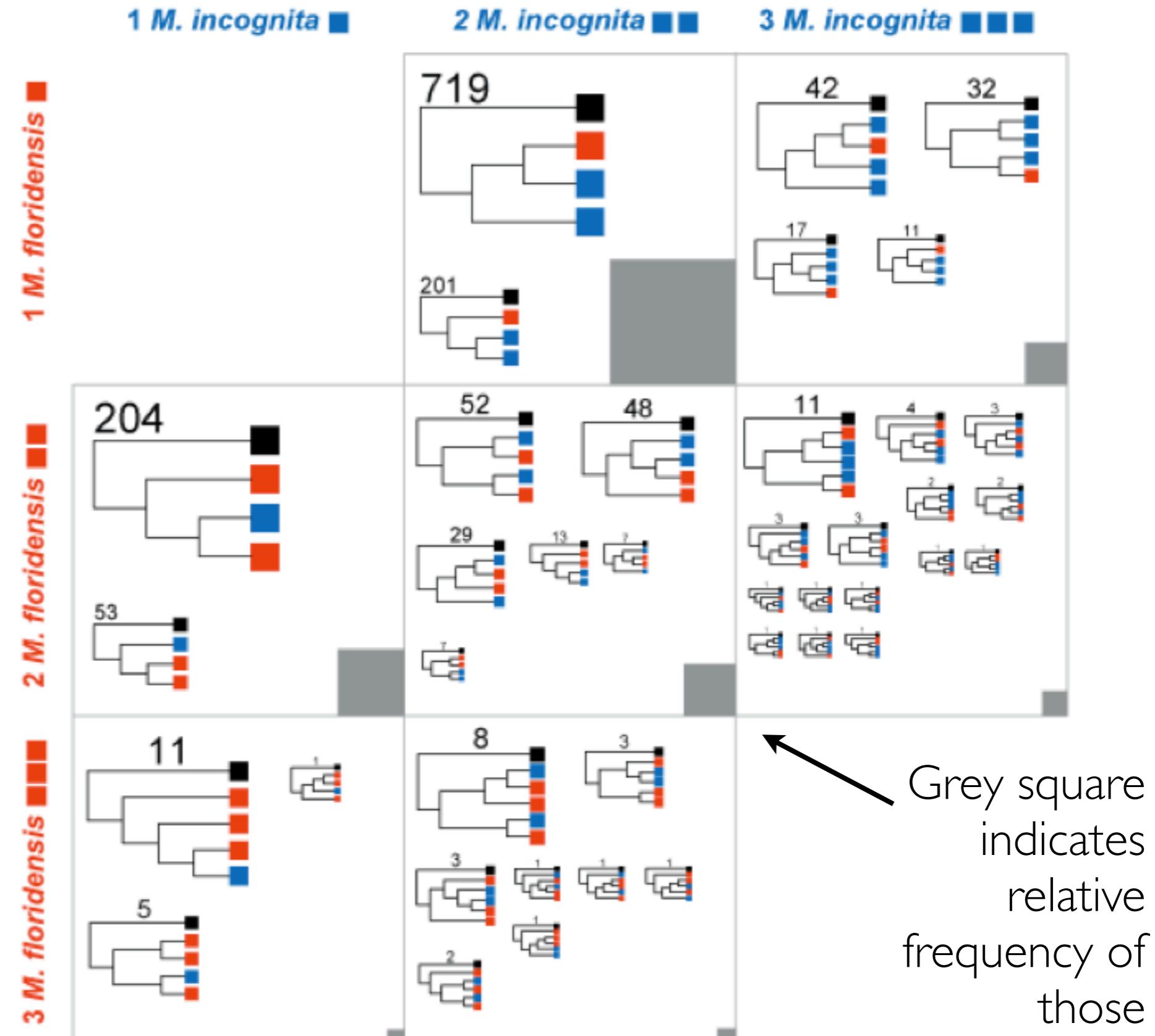
- Coding sequences from 3 genomes were placed into orthologous groups (InParanoid)
- 4018 ortholog clusters included all 3 species
- We retained those with a single copy in the outgroup *M. hapla*
- Phylogenies of relationships between Mi and Mf gene copies (RAxML)
- Trees were parsed and pooled to represent frequencies of different relationships



Each tree contains a single *M. hapla* sequence as outgroup (black square)

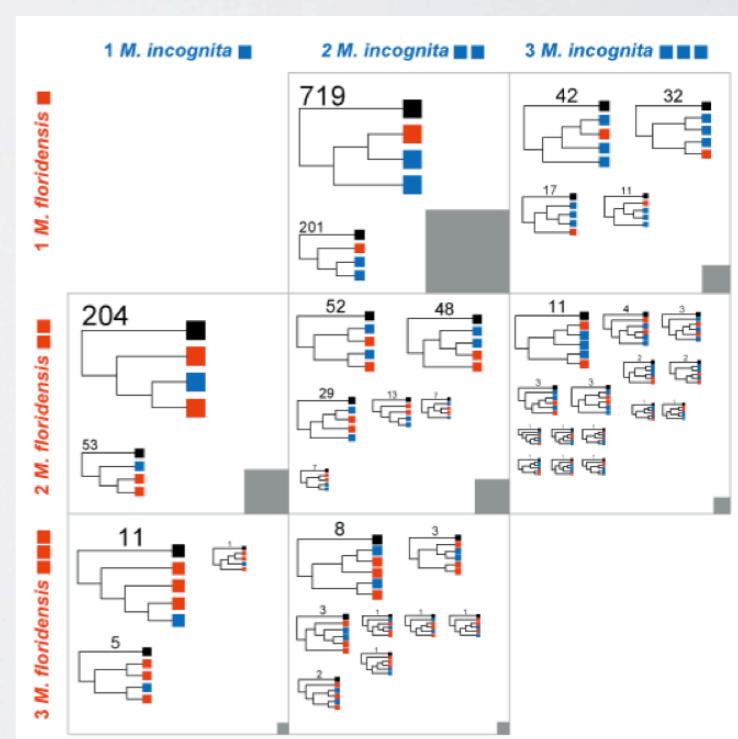
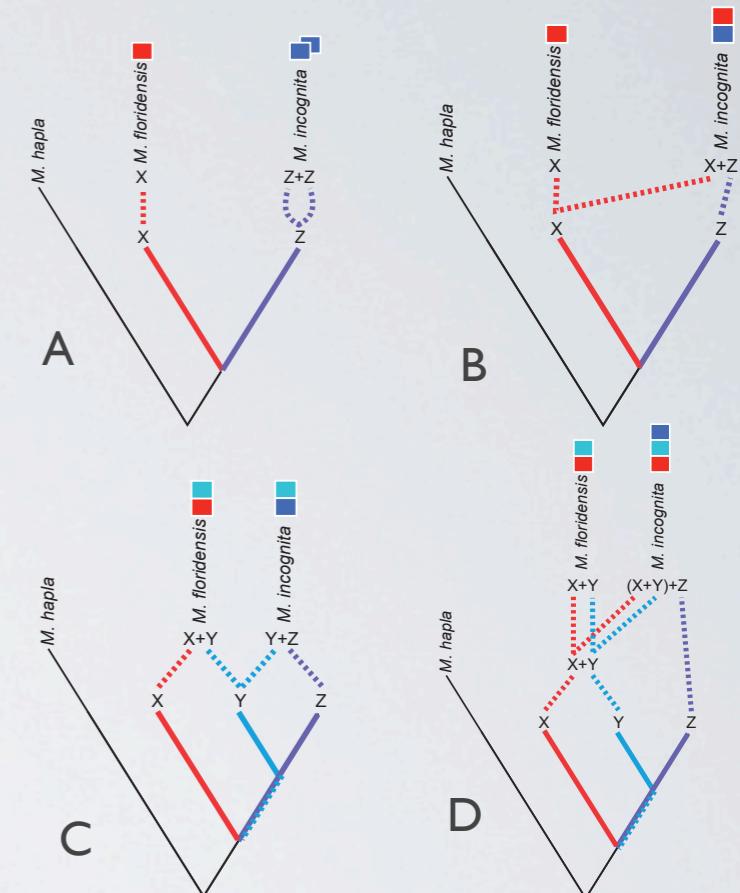
Grid squares represent different numbers of gene copies

Trees are pooled within squares into different patterns of relationships



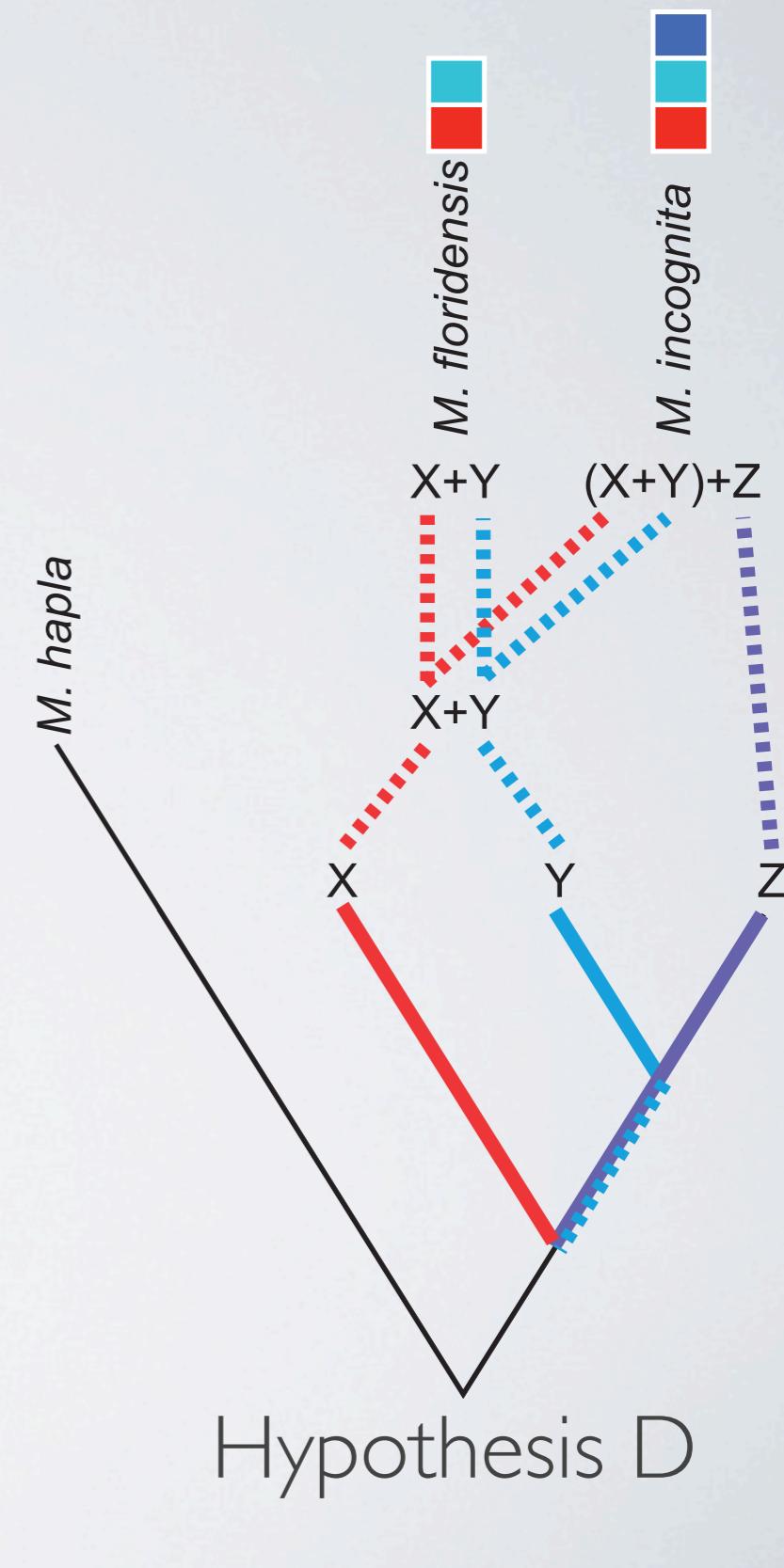
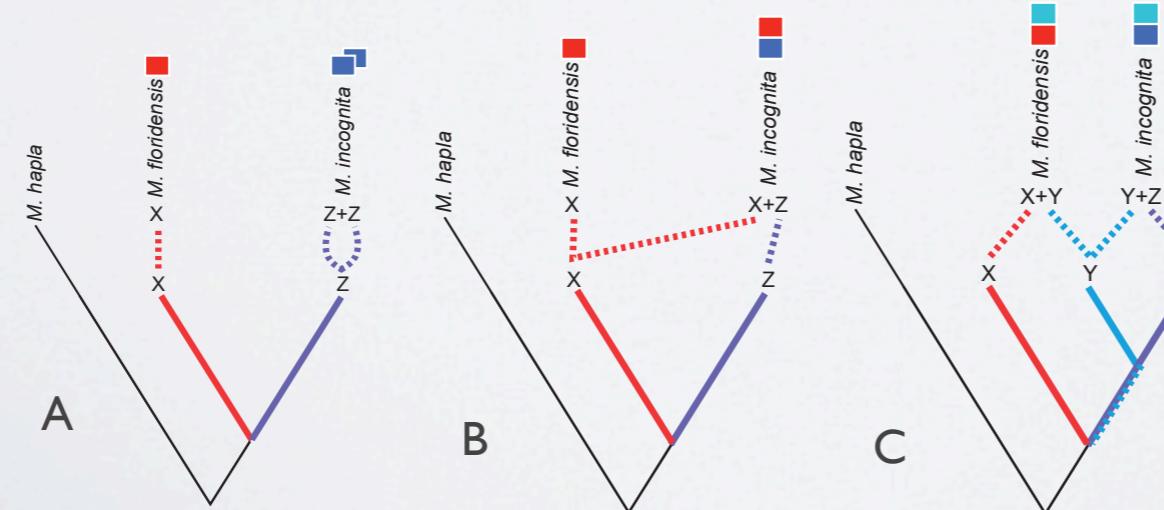
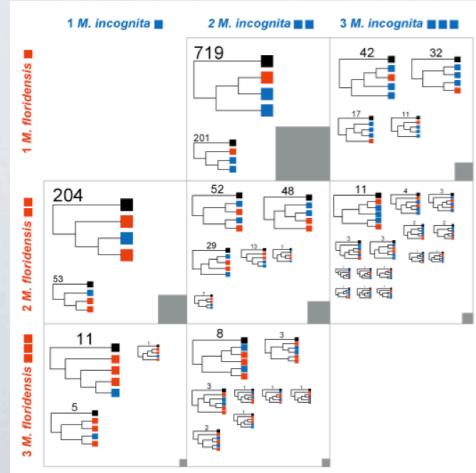
# Testing by Phylogenomics

- We assess the fit of the tree topologies to our hypotheses
- Five out of seven cluster sets, and 95% of all trees, support hybrid origins for both *M. floridensis* and *M. incognita*
  - ie exclude hypotheses A and B
- Hypothesis C best explains 17 trees
- Hypothesis D best explains 1335 trees



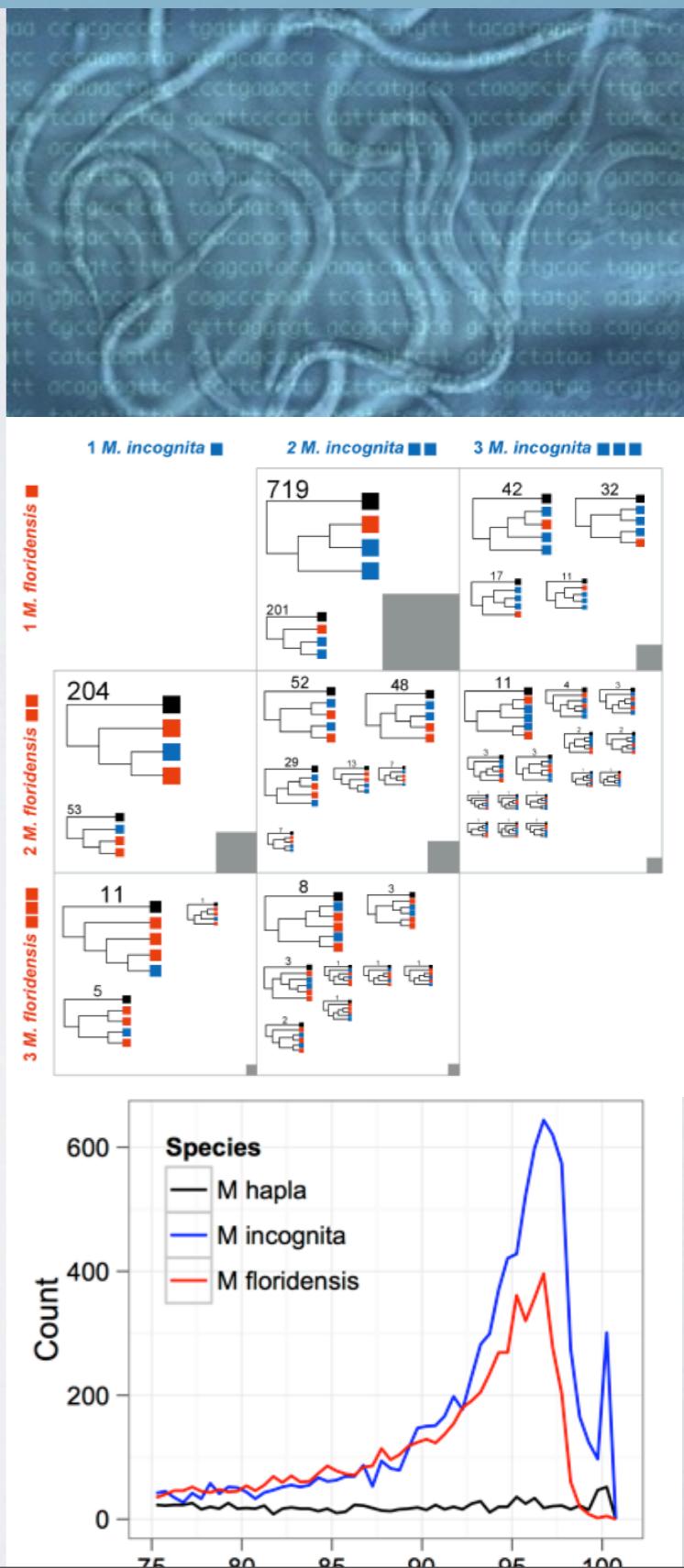
# Testing by Phylogenomics

- The genome data supports both *M. incognita* and *M. floridensis* as interspecific hybrids
- M. floridensis* is a parental species of “double hybrid” *M. incognita* with other parent unknown



# Comparative genomics questions

- Is there evidence of hybrid speciation?
  - Yes, complex hybrid origins are clear
- Is *M. floridensis* a parental?
  - Yes, identified by phylogenomics and allelic sequence identity
- How do offspring and parental genomes differ? What are the broader implications?
  - Ongoing work...

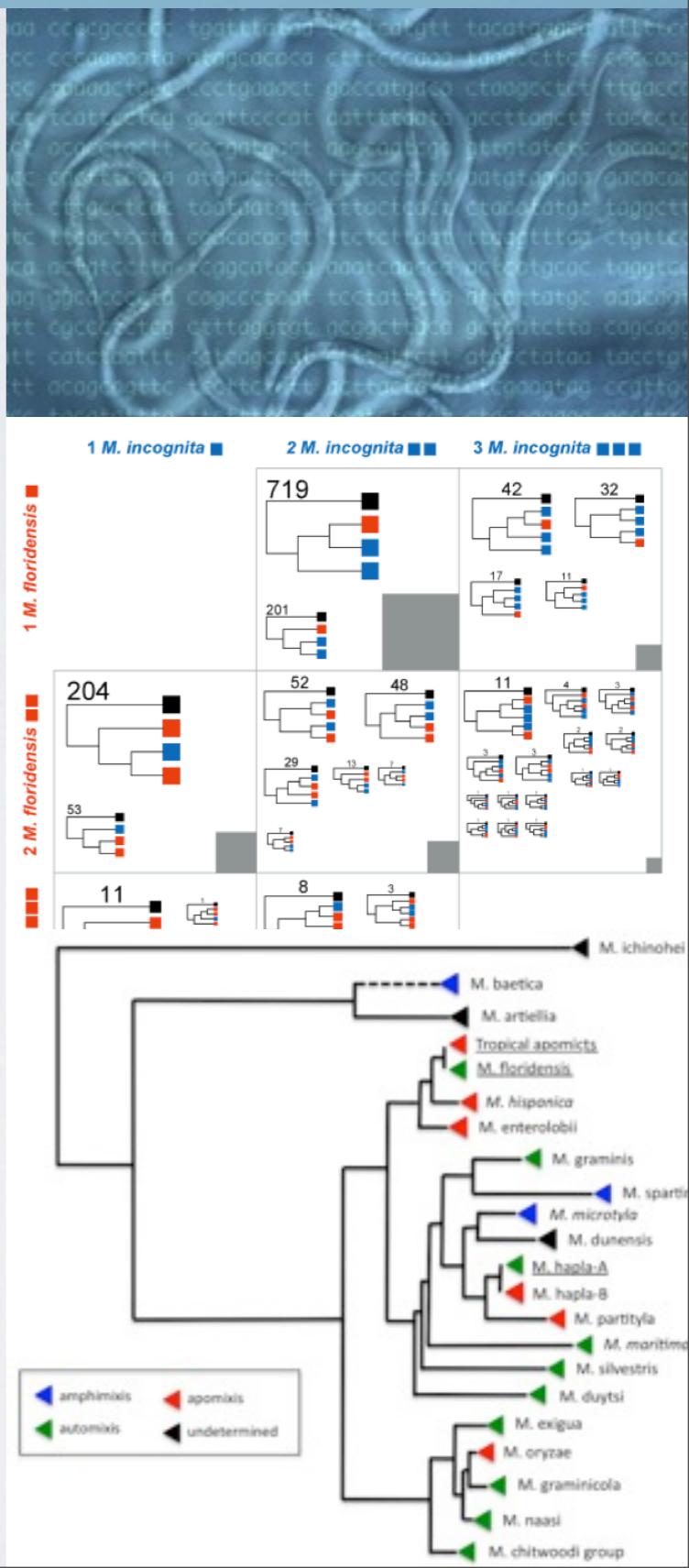


# Ongoing Work

Current NERC grant on Meloidogyne breeding system and genome evolution

- 19 genomes in a phylogenetic design
- Testing effect of recombination & breeding system on genome change
  - hybrids, inbred, outbred, loss of meiosis
  - TEs, mutational patterns, gene families

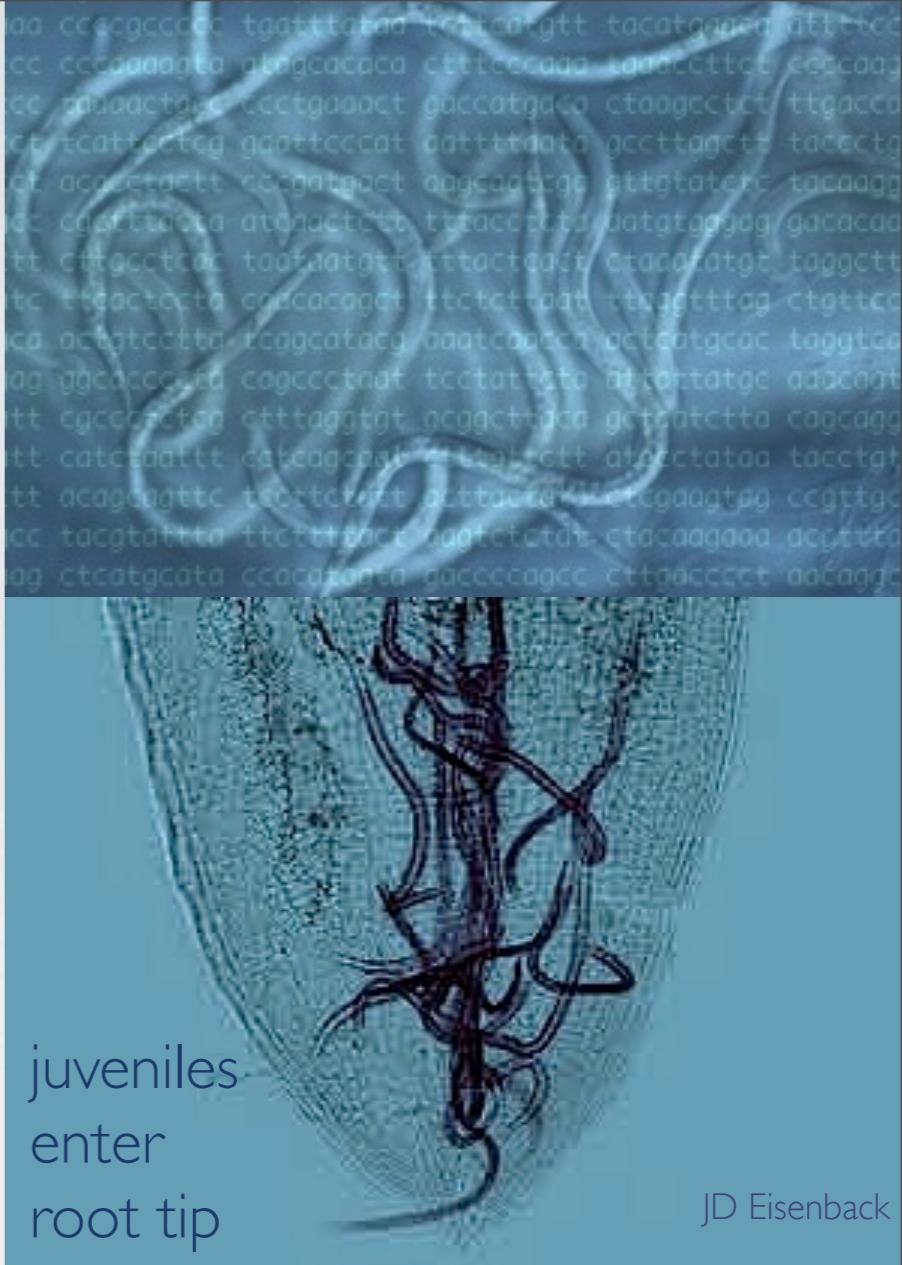
Recombination and genomic rates and patterns of molecular evolution



# COMPLEX HYBRID ORIGINS OF ROOT KNOT NEMATODES

Dave Lunt

Evolutionary Biology Group, University of Hull



# COMPLEX HYBRID ORIGINS OF ROOT KNOT NEMATODES

## Dave Lunt

Evolutionary Biology Group, University of Hull



@EvoHull



+EvoHull



davelunt.net



dave.lunt@gmail.com



@davelunt



+davelunt



<http://www.github.com/davelunt>

