

Species delimitation of a moss clade in a global hotspot for bryophyte diversity

Resolving ambiguities in a clade of *Homalothecium*

Larke E. Reeber



H. megaptilum (bryophyteportal.org)



H. nuttallii (fenzenmosses.com)

Homalothecium characteristics

- Pleurocarpous moss
- Yellow-green, shiny
- On trees, rocks, rocky soil
- Long plicate leaves w/o awns
- Branches curled when dry, spreading when wet



H. pinnatifidum (bryophyteportal.org)



H. fulgens
(blogs.ubc.ca)

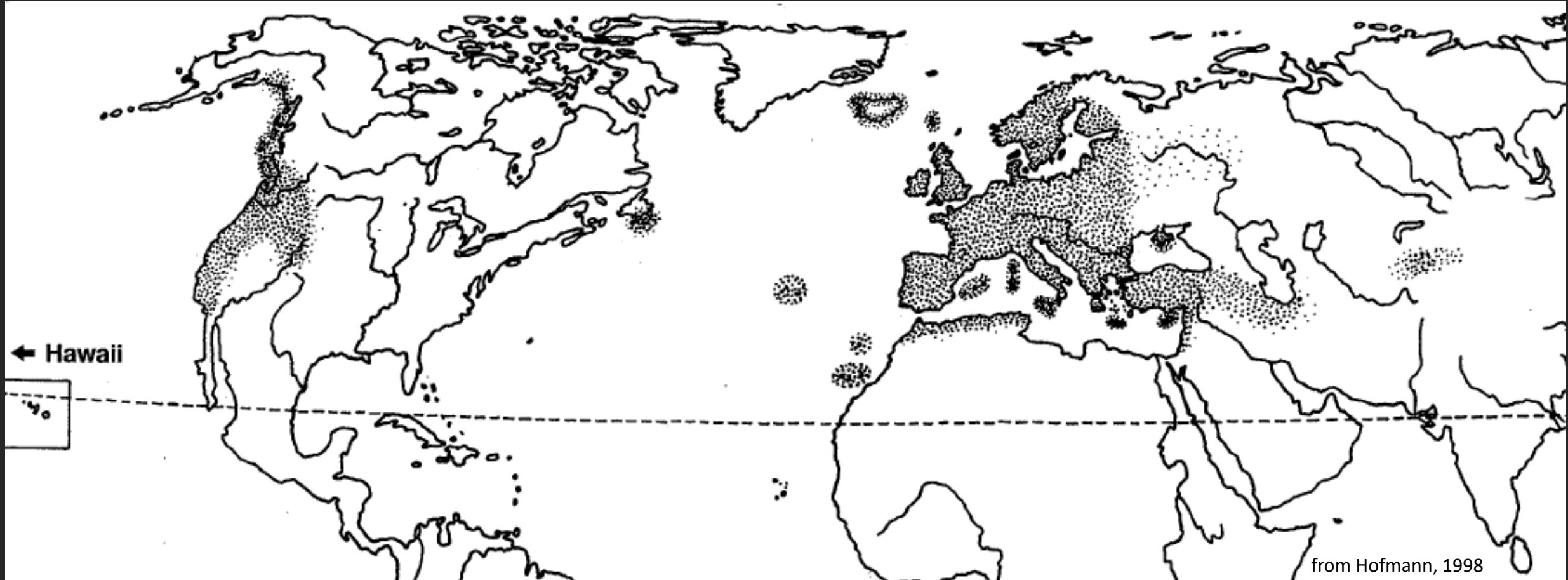


H. aureum
(portugal.inaturalist.org)



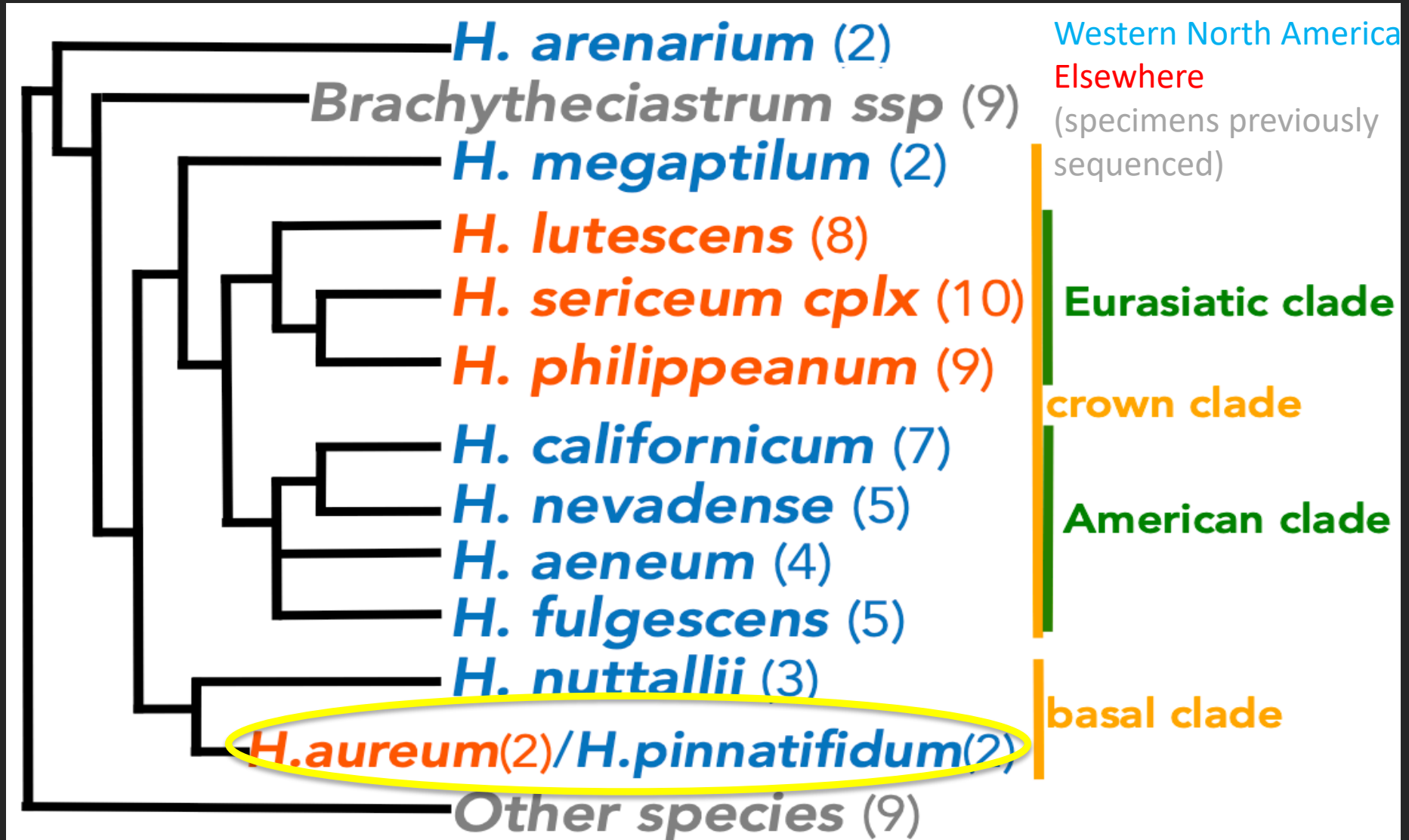
H. sericeum
(eol.com)

Homalothecium--Distribution



Homalothecium

Current Phylogeny



Why this clade?



H. pinnatifidum
(typical)

on rocks/rocky soil

little/no decurrencies



H. pinnatifidum
(atypical)

on leaf litter

long/substantial decurrencies



Also . . .

H. aureum/H. pinnatifidum

- Differing opinions
- Key diagnostic feature descriptions variable
- Number of samples in original molecular study small

Questions

- Is there more than one species within *H. pinnatifidum*?
- What is the *H. aureum* and *H. pinnatifidum* relationship?

Homalothecium--Hypotheses

- Null Hypothesis: There is one species in this clade—*H. aureum*.
- Alternate Hypothesis: There are more than one species in this clade.

Methods Overview

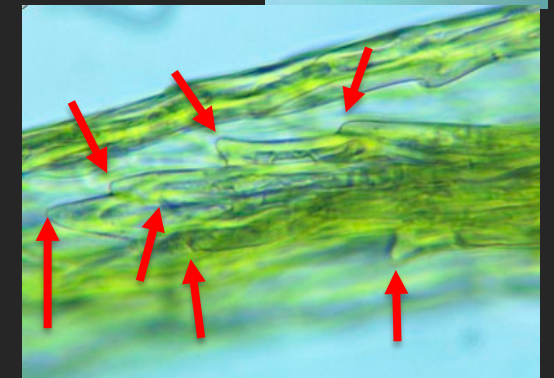
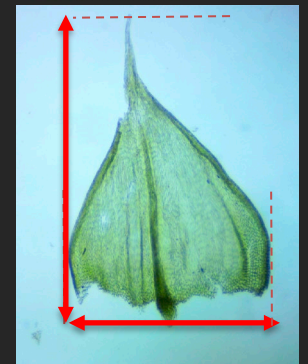
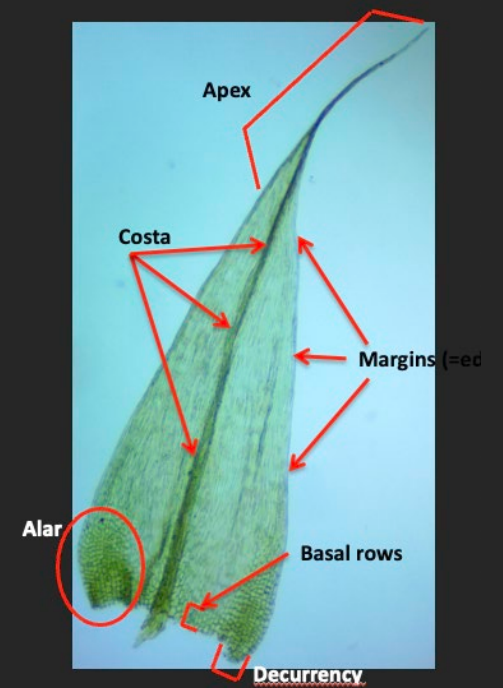
1. Look at LOTS of herbarium specimens of *H. pinnatifidum*.
Define and classify into morphotypes.
2. Select samples for sequencing:
Specimens of each *H. pinnatifidum* morphotype.
Specimens of *H. aureum*
3. Extract DNA and sequence for following gene regions:
ITS1-5.8-ITS2, *atpB-rbcL* and *rpl16* gene regions.
4. Analyze results, incorporating sequences from previous studies.

Morphotyping

- Tracked 11 packet characteristics, 22 gametophytic characters and, when available, 16 sporophytic characters

→ Identified two unique morphotypes that did not fit current *H. pinnatifidum* taxonomic descriptions.

→ Identified several other morphotypes that were intermediate between *H. pinnatifidum* and another *Homalothecium* species.

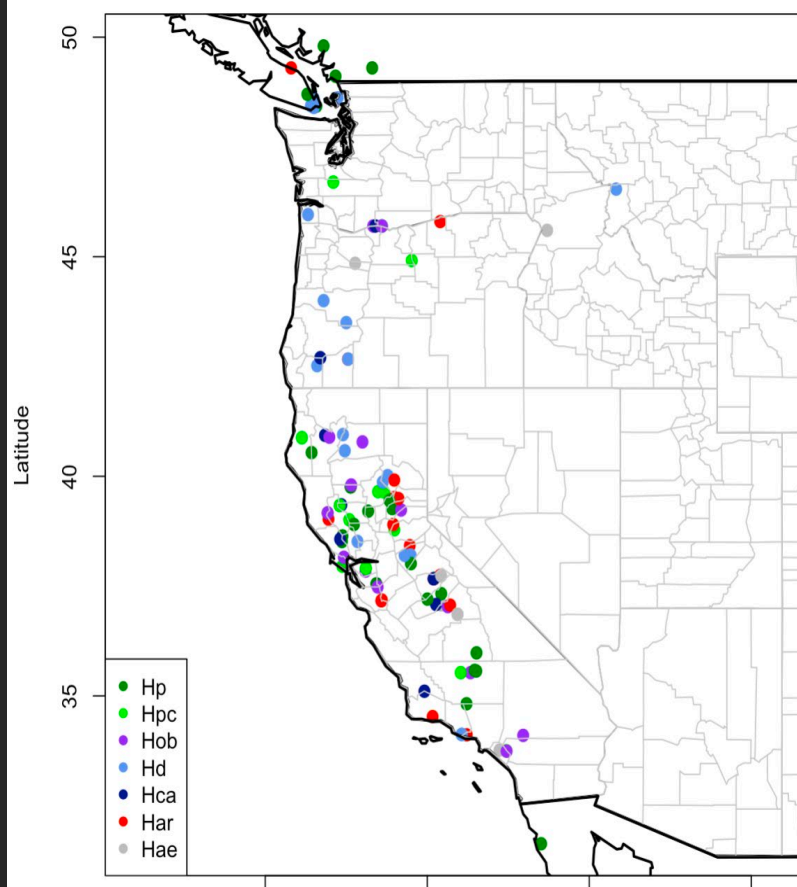


The numbers

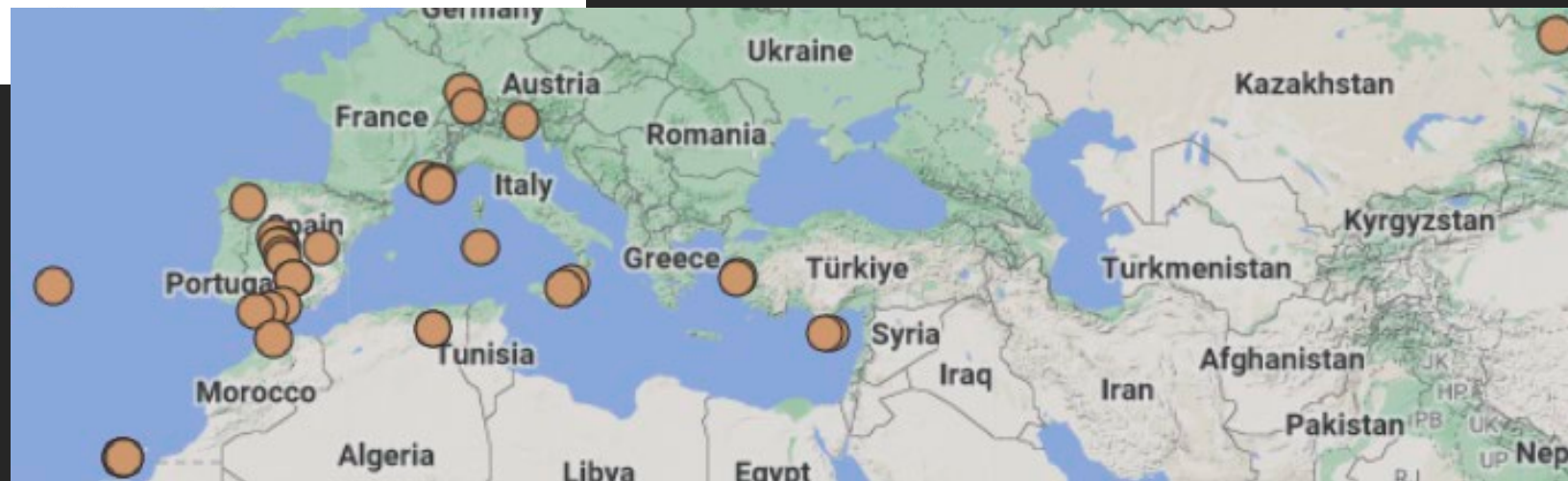
- 446 specimens examined at packet level
- 211 examined microscopically; classified into morphotypes
- 72 specimens had DNA extracted and sequenced
 - 64 were *H. pinnatifidum*:
 - sampled each morphotype across geographic range
 - 12 paired specimens (same collector/ location/date, different morphotype)
 - 7 were *H. aureum*: across geographic range;
 - 1 was *H. nuttallii*

Study samples

- REDO WNA MAP;
- ADD CORRECT *H. AUREUM* MAP



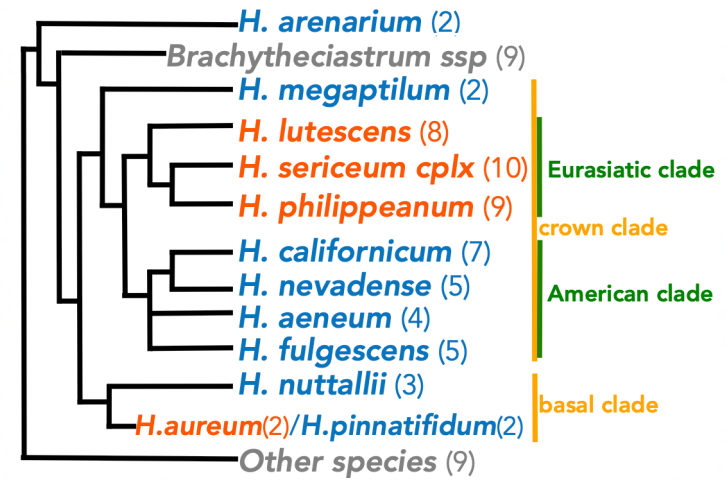
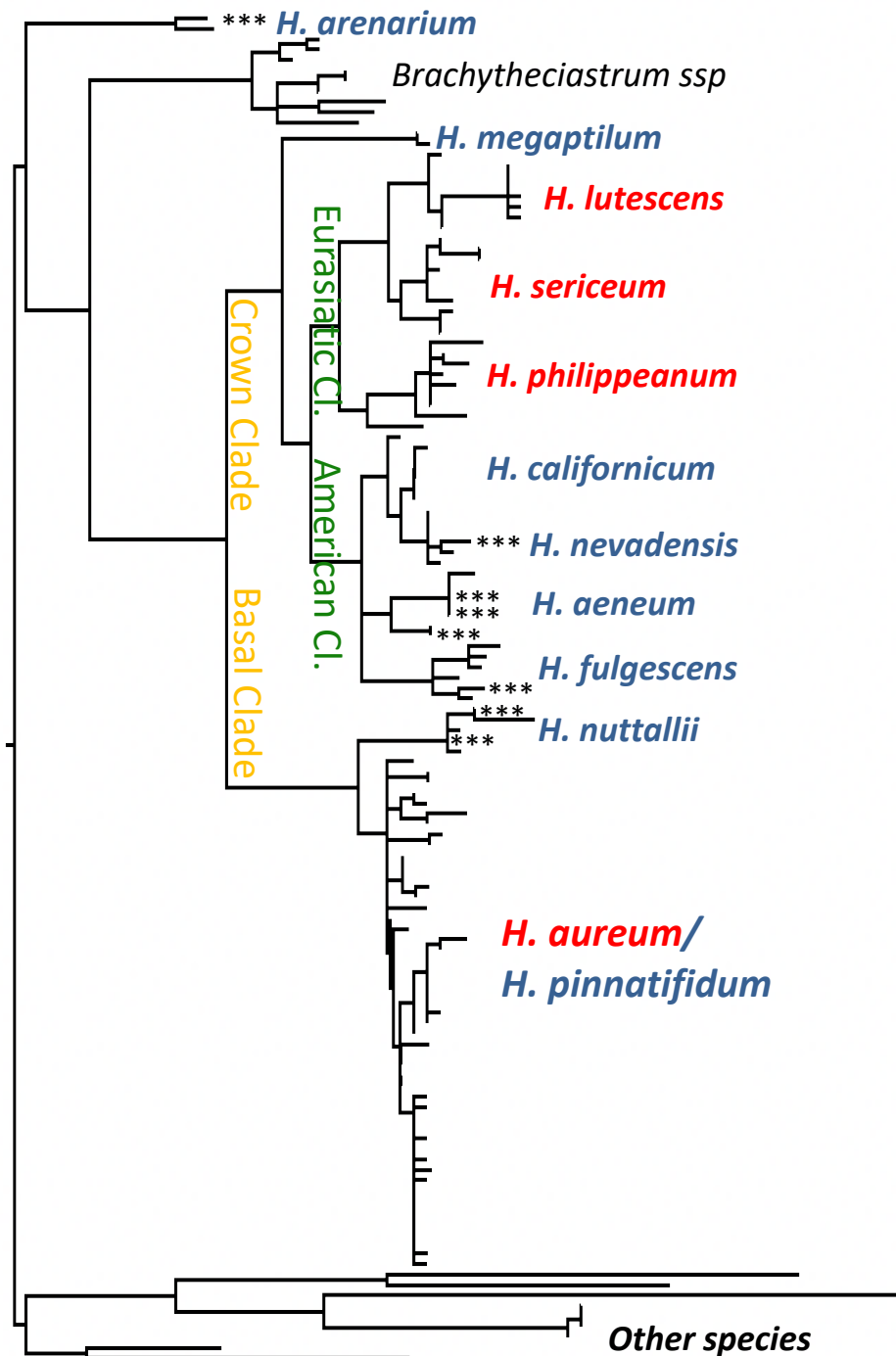
from Hofmann, 1998



Genetic Analysis--Methodology

- All four genes concatenated—only specimens with >1 viable gene sequence
- Indels encoded per simple encoding method of Simons and Ochoterena (2000)
- Tree generated by RaxmL (ML, 100 bootstraps) with separate rate partitions for chloroplast (rpl16, atpB_rbcL) (GTR) ,nuclear genes (ITS1, ITS2)(GTR), and indels (BIN)
- FigTree used to visualize the trees

Gene region	rpl16	atpB-rbcL	ITS1	ITS2
Viable sequences	56	59	43	45
# of BP	730-779	399-713	237-405	421-437
Trimmed size				
Non-unique BP	Add in trimmed size used, # of non-unique bp, # of phylogenetically useful bp, # of phylogenetically useful indels (should I do this, are these right descriptions, what should really be included?)			
Phylogenetically useful BP				
Phylogenetically useful indels				



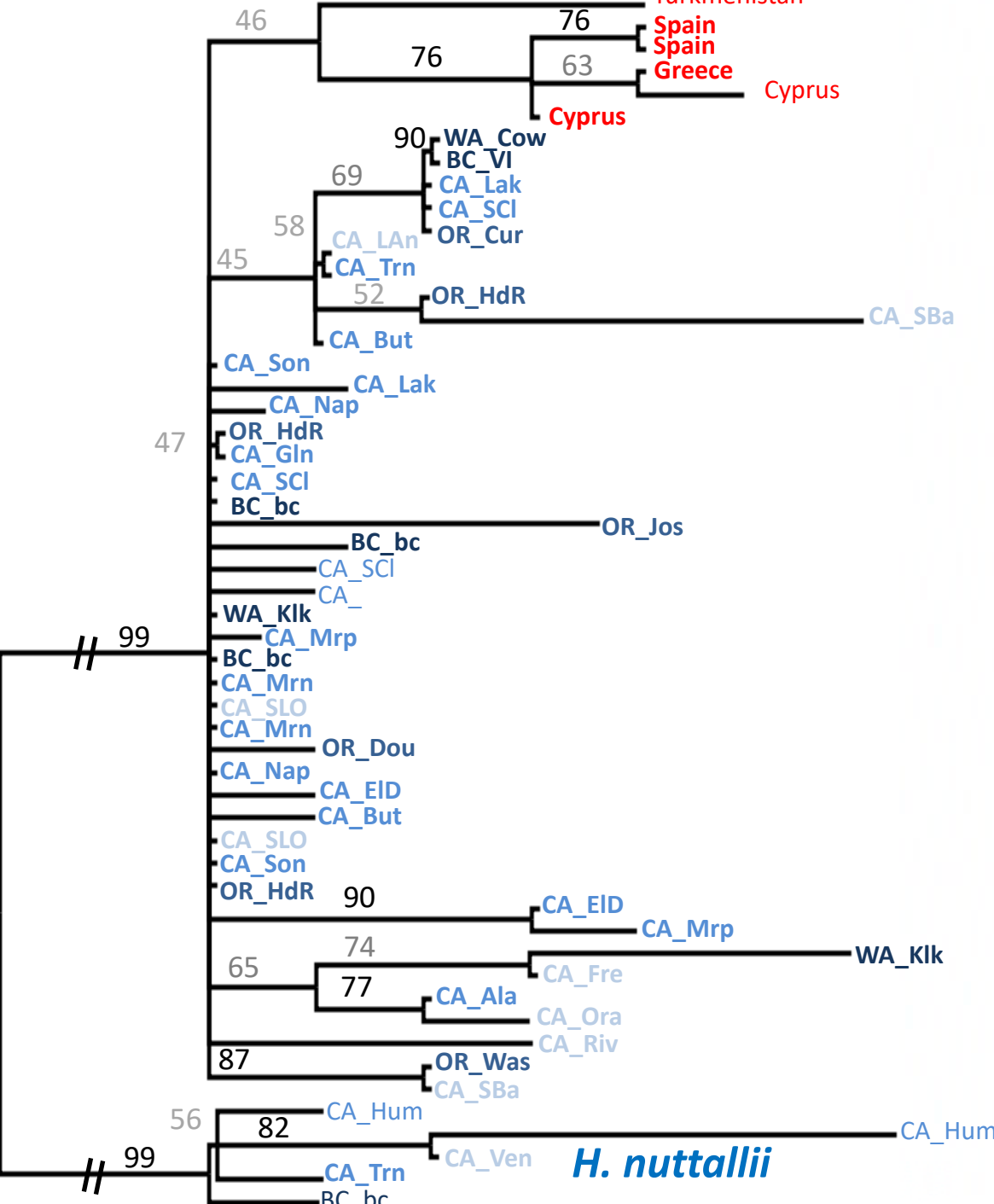
All sequences with >1 viable gene sequence

→ Recovered nearly identical overall phylogeny as previous papers, including basal clade with *H. nuttallii* sister to a *H. aureum*/*H. pinnatifidum* clade

→ Some (8), but not all, of the "intermediate" morphotypes specimens are a better genetic match to a different *Homalothecium* species

Basal Clade only

British Columbia, Washington, Montana
 Oregon, Idaho
 California—northern
 California—southern
 >75
 60-75
 45-59



- *H. nuttallii* well supported monophyletic clade
- Eurasian sequences (ie; *H. aureum*) form a very weak monophyletic clade
- Western North American sequences (ie; *H. pinnatifidum*) do not form a clear monophyletic clade separate from Eurasian sequences

Implications

- No clade of any morphotype within *H. pinnatifidum*
- *H. aureum* in a clade but incompletely differentiated from *H. pinnatifidum*

Acknowledgements

- Funding:
 - Northern California Botanical Society
 - San Jose University Biological Sciences Dept.
 - Anonymous funder
- Advising:
 - Dr. Benjamin Carter
 - Dr. Tracy Misiewicz
 - Dr. Susan Lambrecht
 - James Shevock
- Herbaria Specimens from: CASC, UBC, SJSU, MO, CHSC
- General support: Jack Fendell, John McLaughlin, Charlotte Bony, Charlotte Miranda, Andy Frank



Add in herbaria logos

Basal clade only

Homalothecium Questions??

Homalothecium
Questions??

Homalothecium

Backup

Preliminary
Results
--Haplo Networks

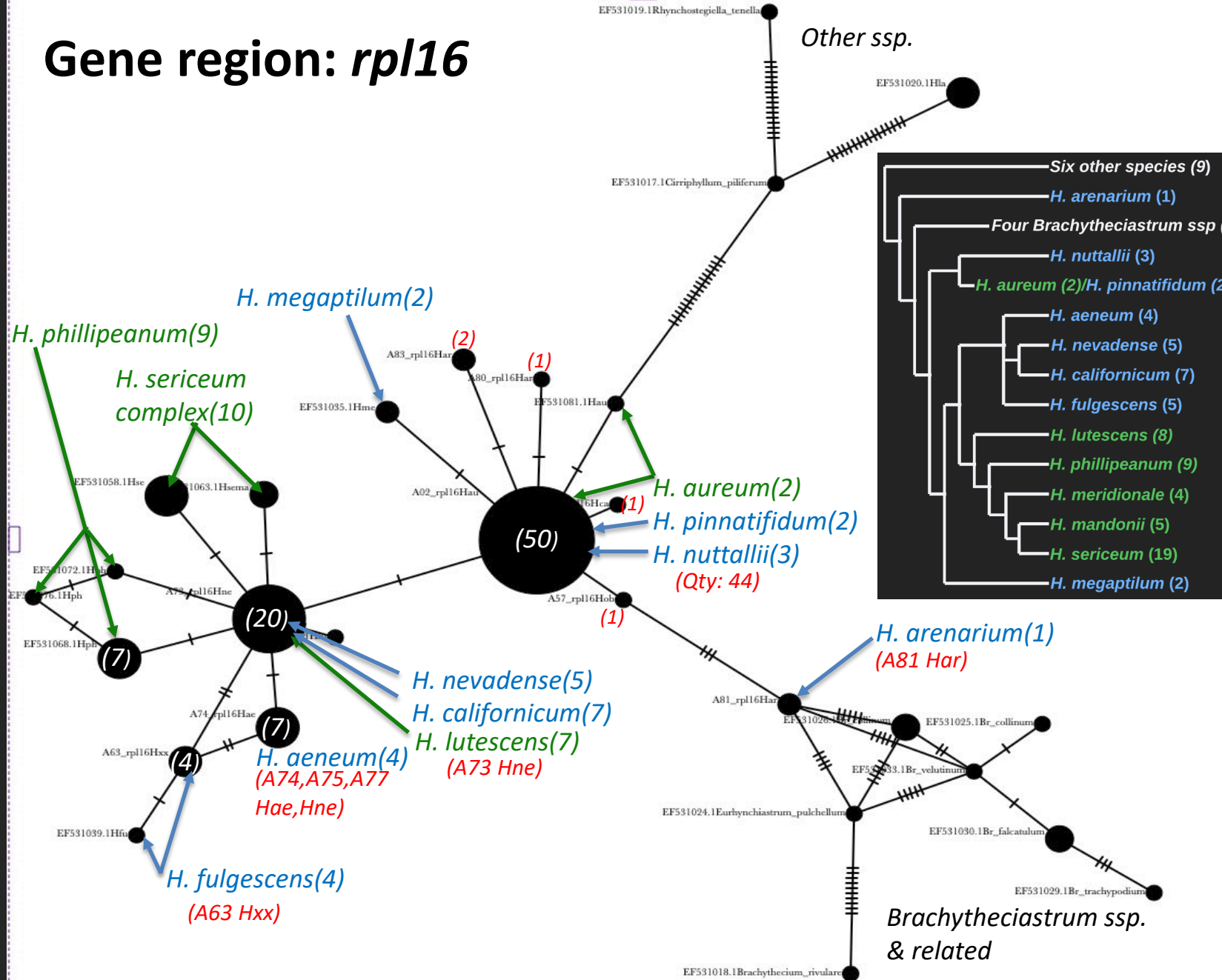
Methodology

- Genious 2023.2.1: Build contigs, build and trim alignments, create Nexus output
- PopArt (2023): Build Haplone networks
- Consensus sequence built from forward and reverse sequence of each specimen
 - Consensus sequence used for alignment (default settings)
 - If no consensus sequence, used higher quality of forward/reverse sequences
- Alignment building (for each gene) (default settings)
 - Start with consensus or best quality sequence for each specimen
 - After initial alignment, removed grossly misaligned sequences
 - Remove very poor quality/very incomplete sequences
 - Trimmed sequence ends so that most sequences were complete end to end
 - Exported as Nexus alignment
- Built Haplo network (PopArt)
 - Import Nexus alignment
 - Built Haplone network

Network notes for following pages

- Size of dot is number of specimens that had identical sequences
- Blue/Green names are GenBank accession species in that dot
- Red are *H. pinnatifidum* and *H. aureum* specimens that I sequenced that are in that dot(s) (Axx=specific specimen number, Hxx=morphotype)
- Number in parenthesis are quantities (in dot, of GB accessions, of specimens)
- Networks created in Popart
 - Hashmarks are number of nucleotide differences
 - Indels are not included as a difference
 - Nucleotides with ambiguity in any specimen are excluded in determination
- GenBank accessions: A few sequences were missing from GenBank (i.e.; the accession numbers of some phylogenetic tips were not listed in the paper); Affects *H. sericeum* complex
- “*H. sericeum* complex” includes *H. sericeum s.str.*, *H. mandonii*, and *H. meridionales*

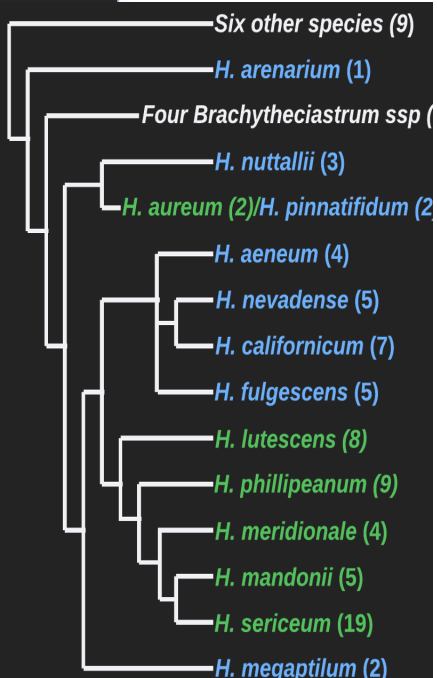
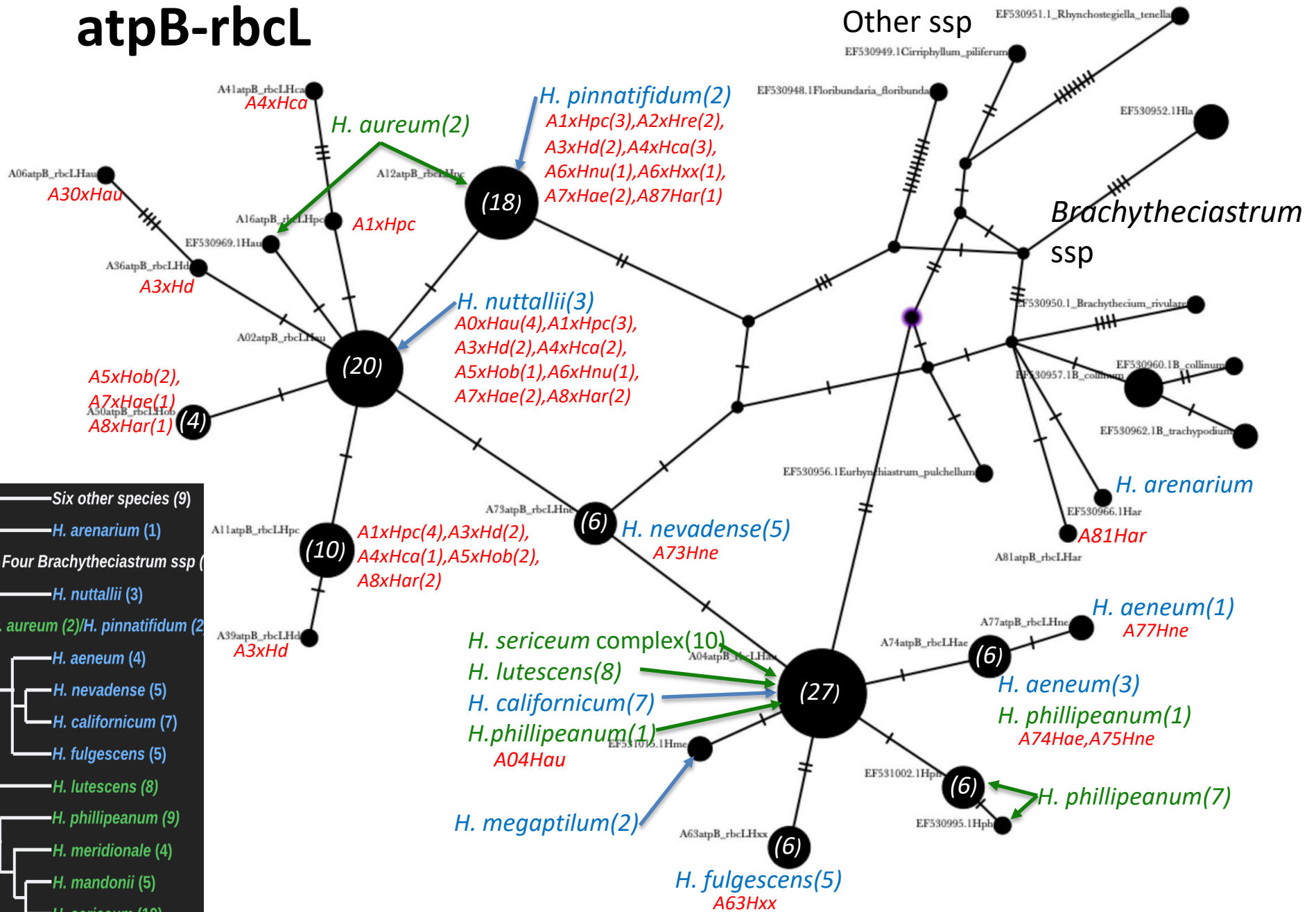
Gene region: *rpl16*



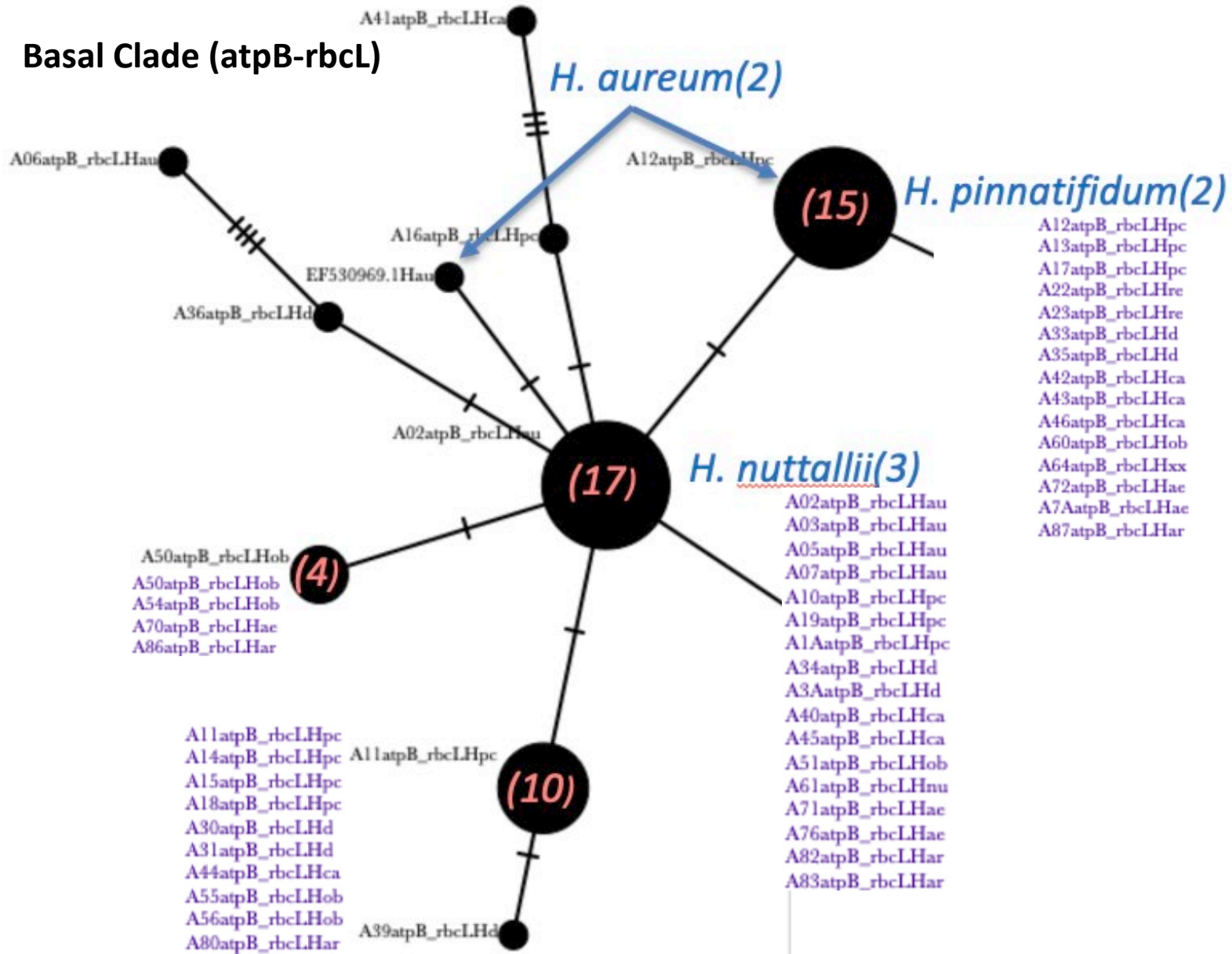
Implications

- I still need to improve my identification skills.
 - A63, A73, A74, A75, A77, A81 probably misidentified
- Only differentiation between *H. nuttallii* and *H. pinnatifidum/H. aureum* (an indel) is not picked up by PopArt
- *None of the morphotypes are differentiated with rpl 16*
- → *No surprises*

atpB-rbcL



Basal Clade (atpB-rbcL)



Implications

- A63, A73, A74, A75, A77, A81 misidentified
- *atpB_rbcL* does differentiate between *H. pinnatifidum*/*H. aureum* and *H. nuttallii*
- None of the morphotypes correspond to the network nodes
- But something may be going on. . .

Investigation

→ A microscopic exam of samples within each group yielded no identifiable distinguishing feature that matched groupings.

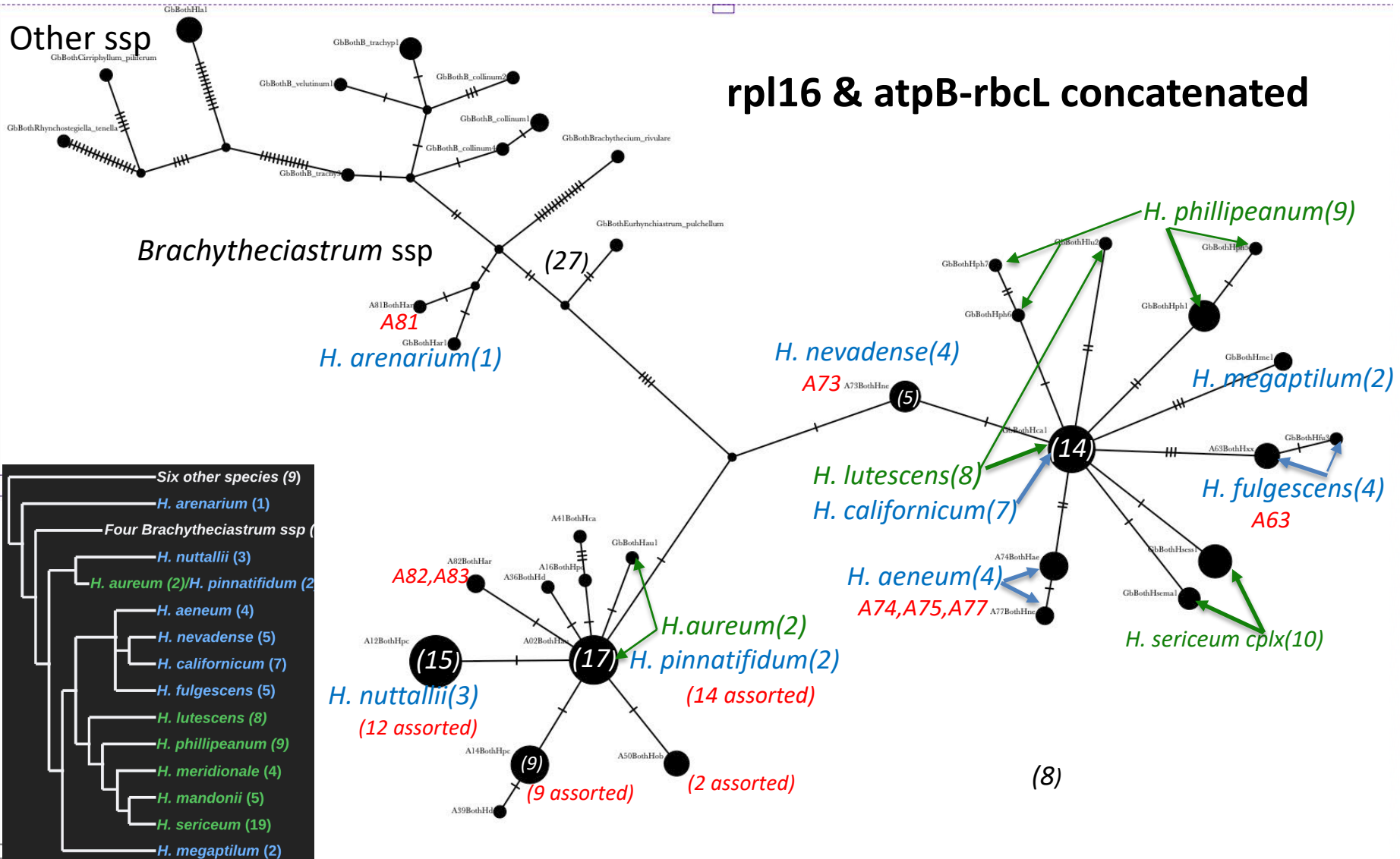
Investigation

→ Two nucleotides (positions 282/386) drove the groupings.
Simple chi-squared test indicated these results may not be significant. [Did I do this right?]

Combination	All	GT	GC	AT	p-value
Sequences	59	29	17	13	
Proportion	1.00	0.491525	0.288136	0.220339	
Hpi (Prop.)	47	18 (0.383)	16 (0.340)	13 (0.277)	0.3250
Hau (Prop.)	7	6 (0.857)	1 (0.143)	0 (0.0)	0.1381
Hnu (Prop.)	5	5 (1.00)	0 (0.0)	0 (0.0)	0.0753

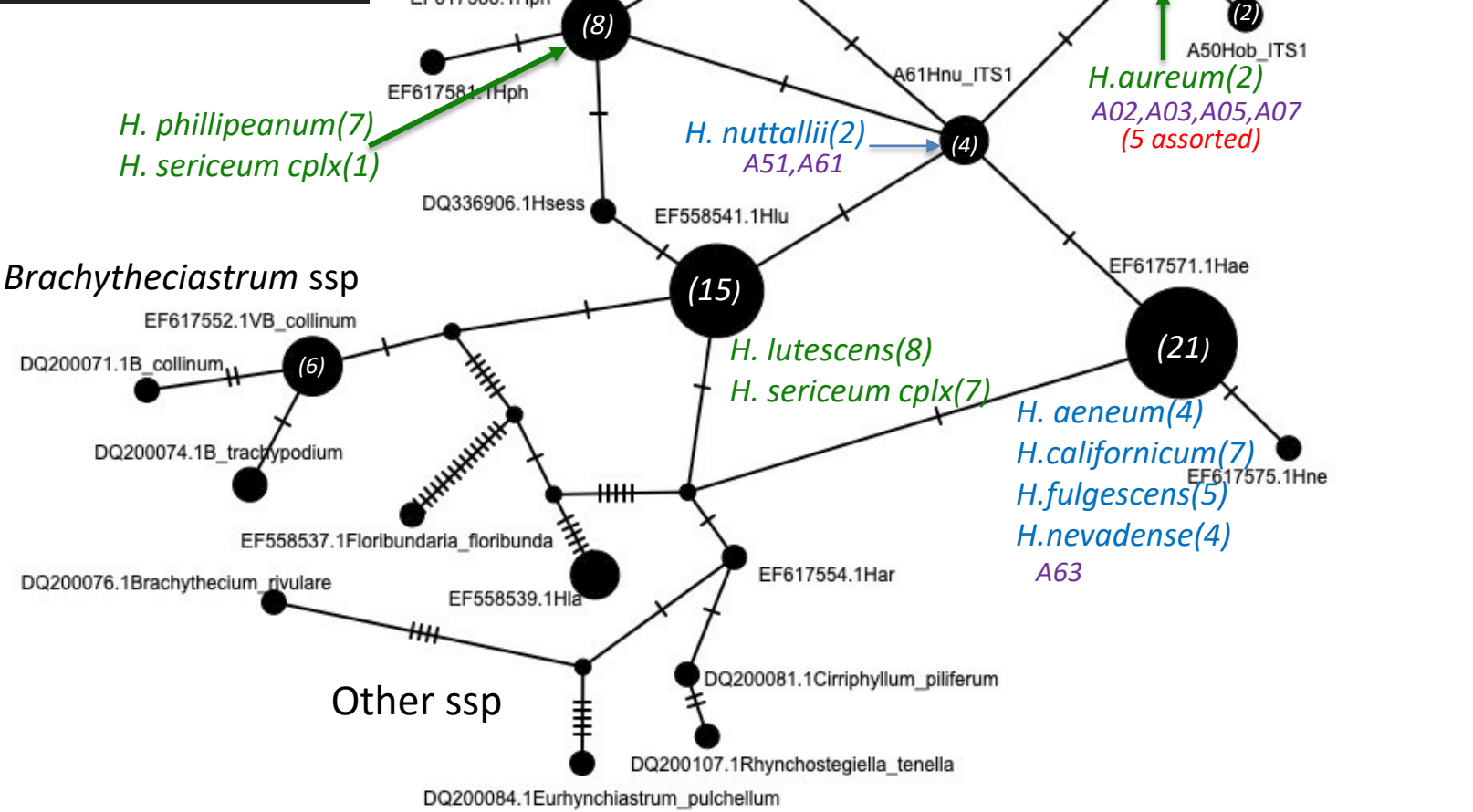
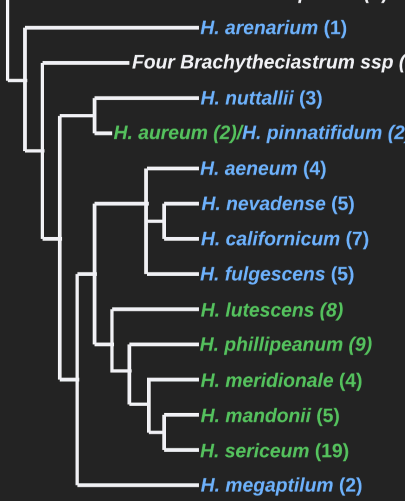
The chi-square approximations may be incorrect because of the small sample sizes for *H. aureum* and *H. nuttallii*, but the preliminary conclusion is that these differences are not statistically significant.

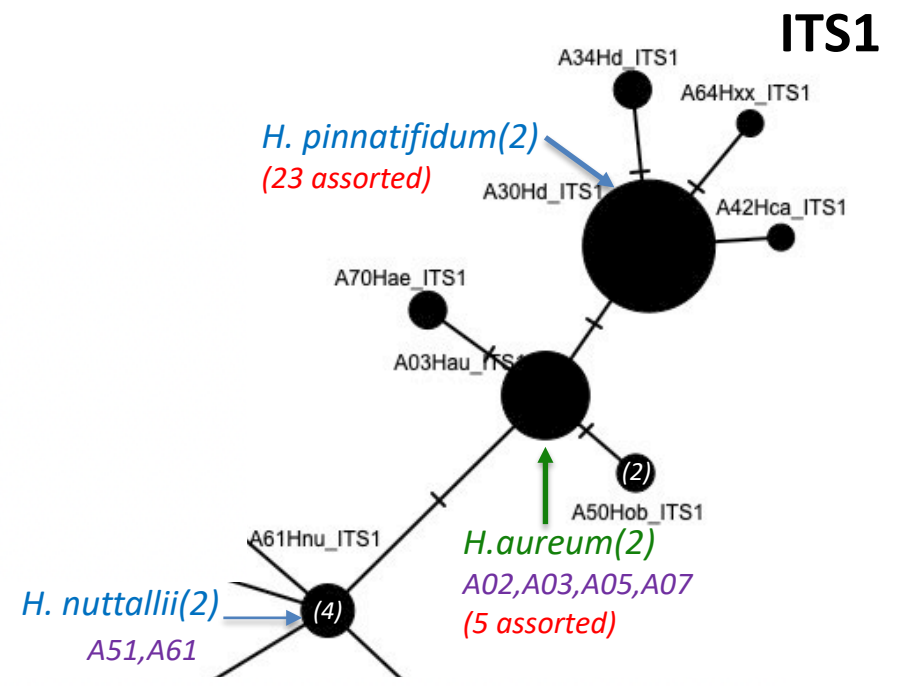
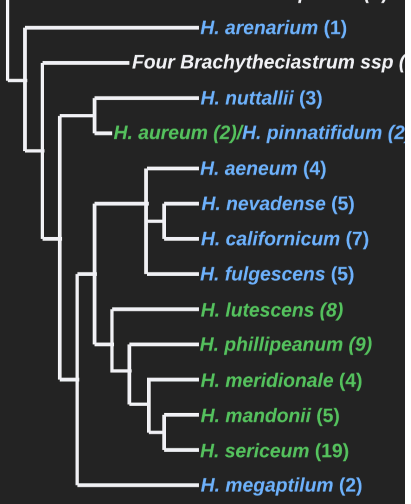
rpl16 & atpB-rbcL concatenated



Implications

- *H. aureum/H. pinnatifidum/H. nuttallii* nodes are essentially unchanged
 - ➔ Likely due to the PopArt not using any indels or nucleotides with any ambiguity codes

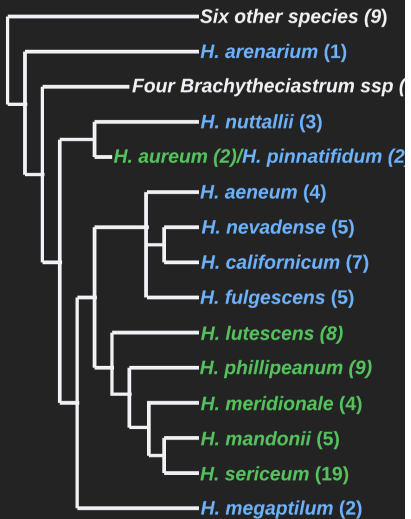
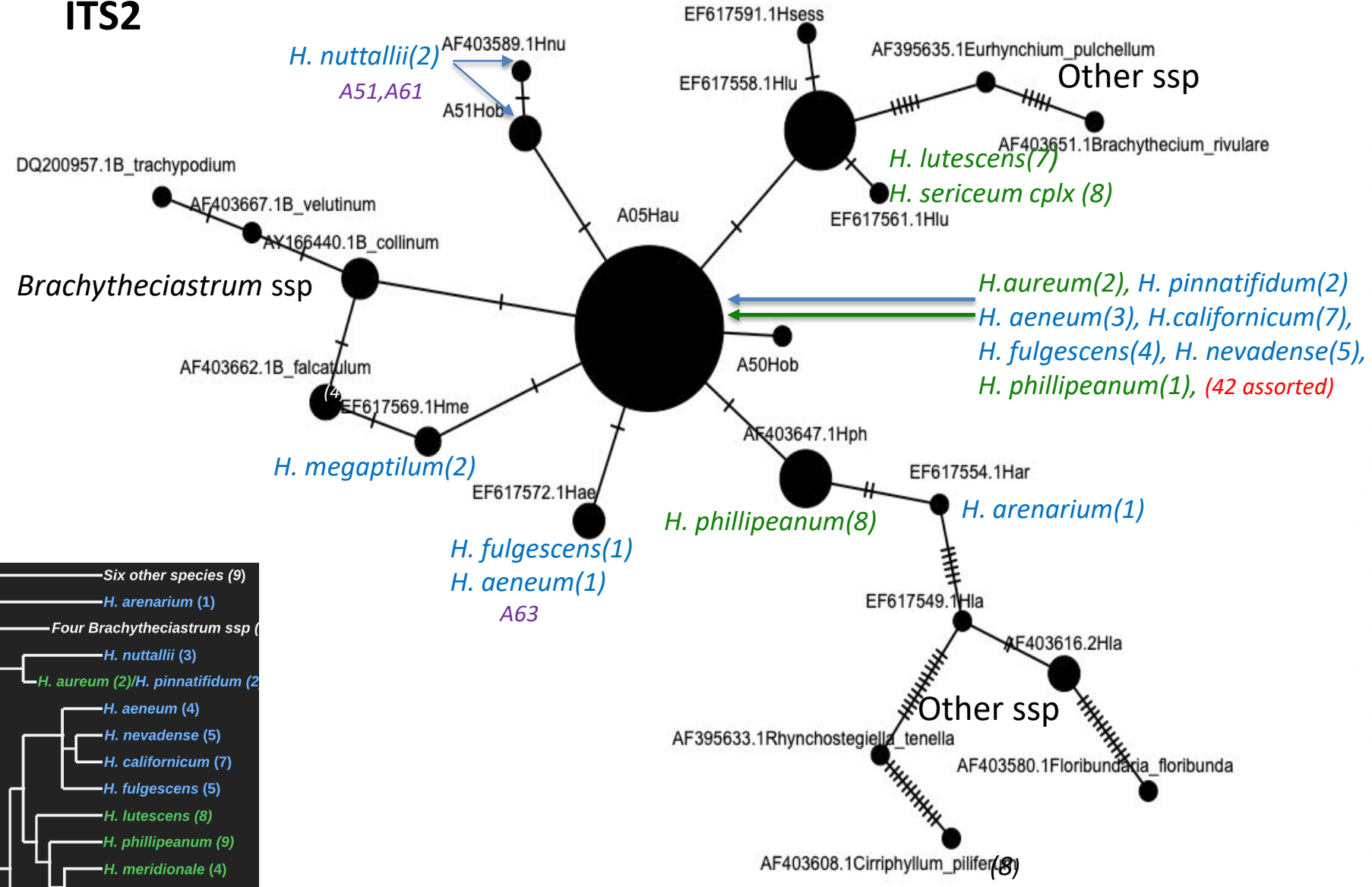




Implications

- *H. nuttallii* separate from *H. pinnatifidum*/
H. aureum
- A51, A61 also misidentified
- *H. pinnatifidum* incompletely separated
from *H. aureum* (all *H. aureum* together,
but also includes some *H. pinnatifidum*)

ITS2



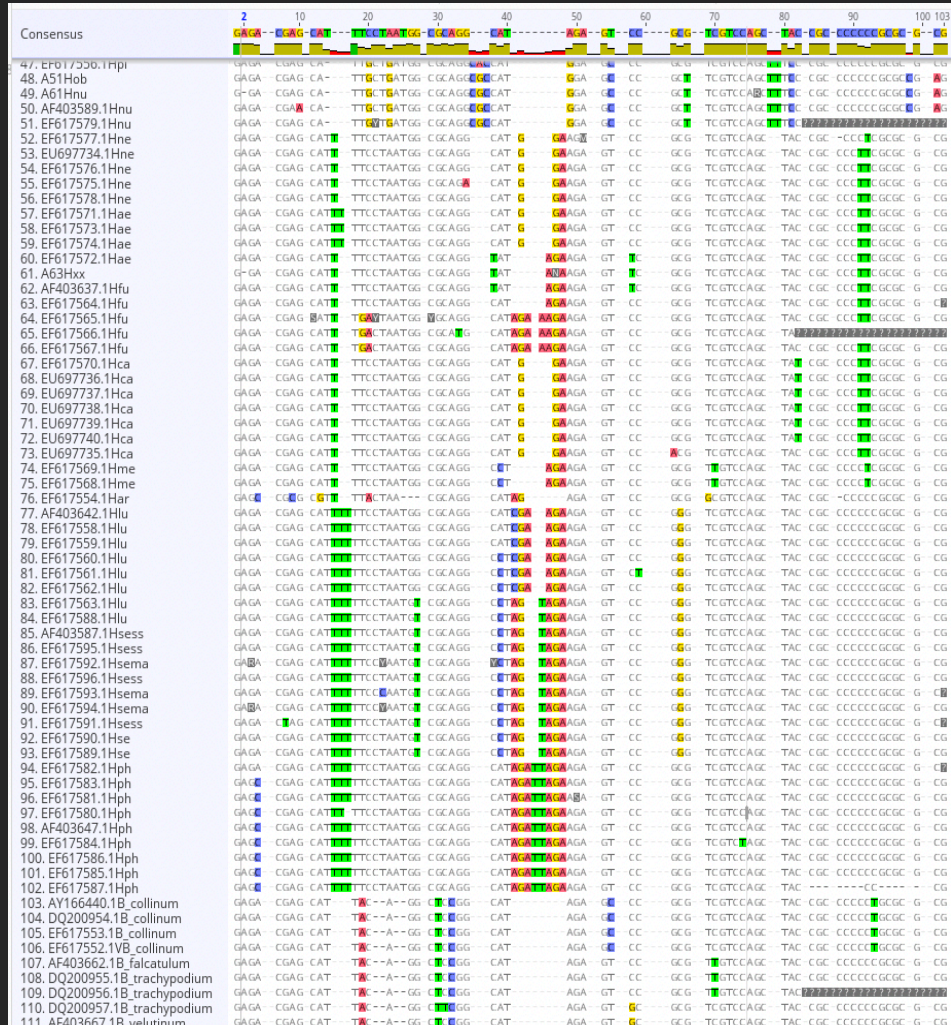
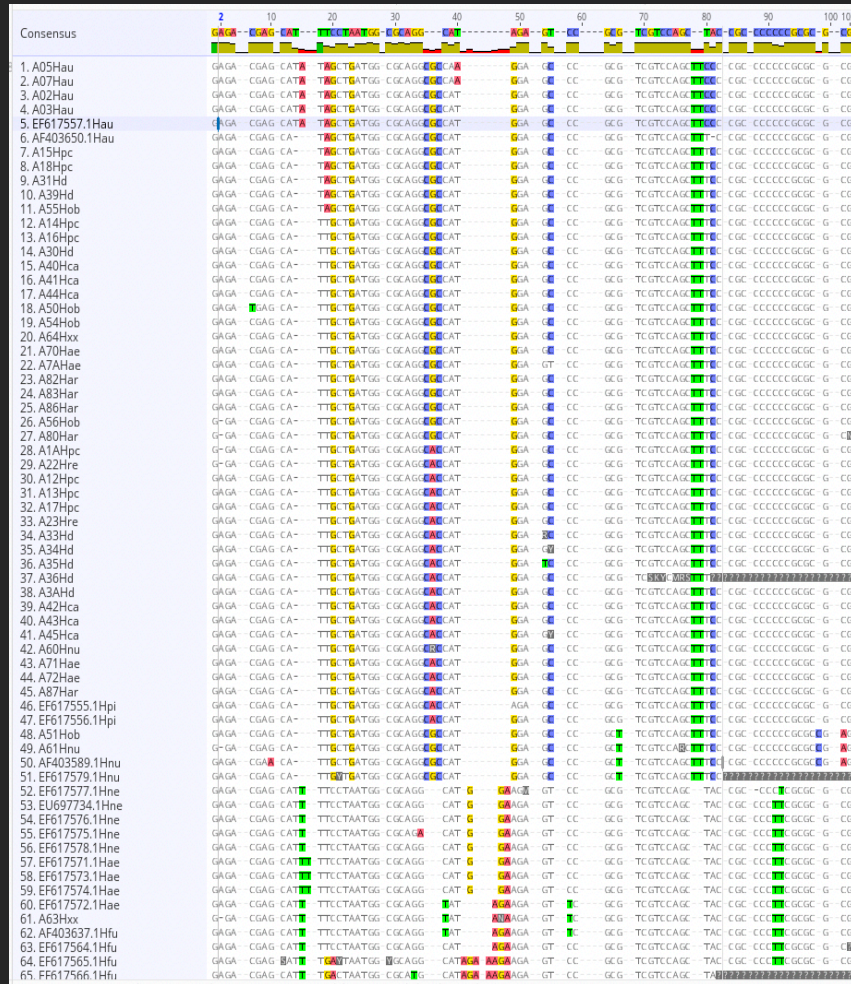
Implications

- *H. nuttallii* separate from *H. pinnatifidum*/
H. aureum
- *H. aureum*/*H. pinnatifidum* not even
completely separated from most other
Homalothecium ssp

→ Why?

Why a big blob?

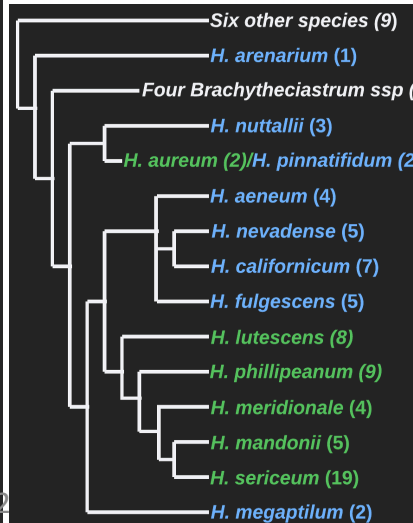
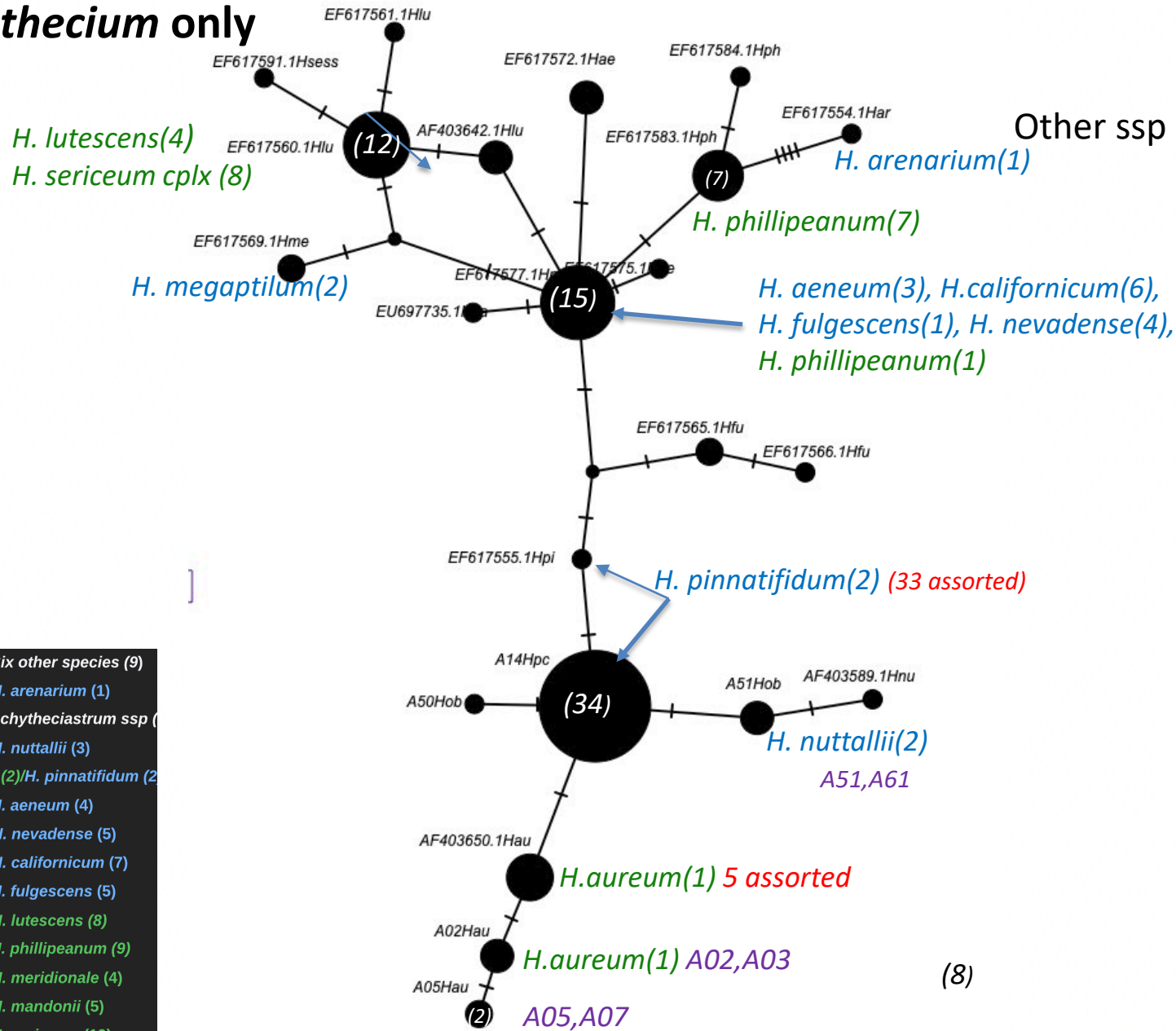
→ Due to non-use of indels and nucleotides with ambiguity codes



ITS2 sequences with most if undifferentiated nucleotides removed
→ almost none of the differences visible above are used

ITS2

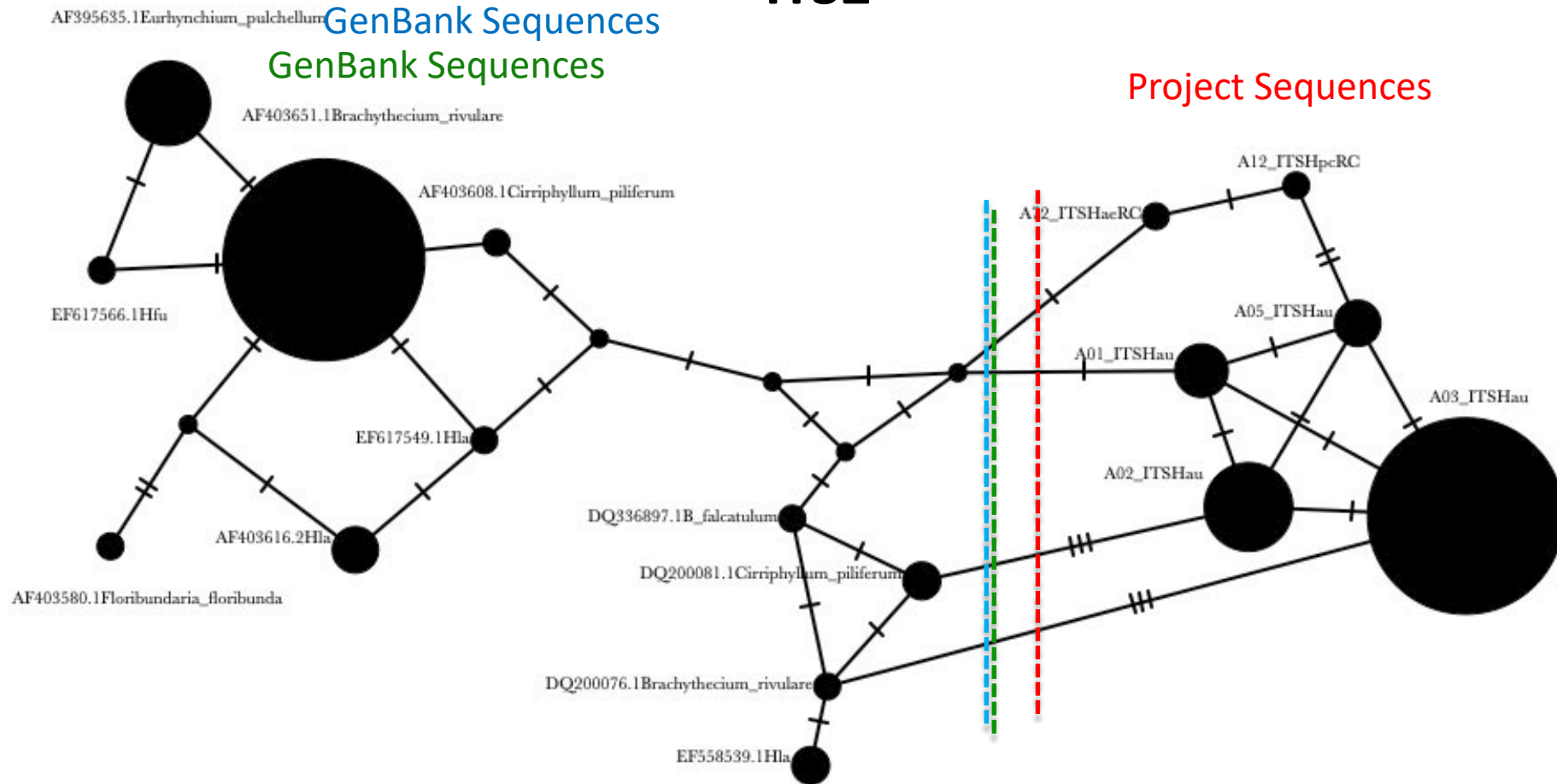
Homolathecium only



Implications

→ The avoidance of indels/ambiguities are making the use of the PopArt tool not terribly useful

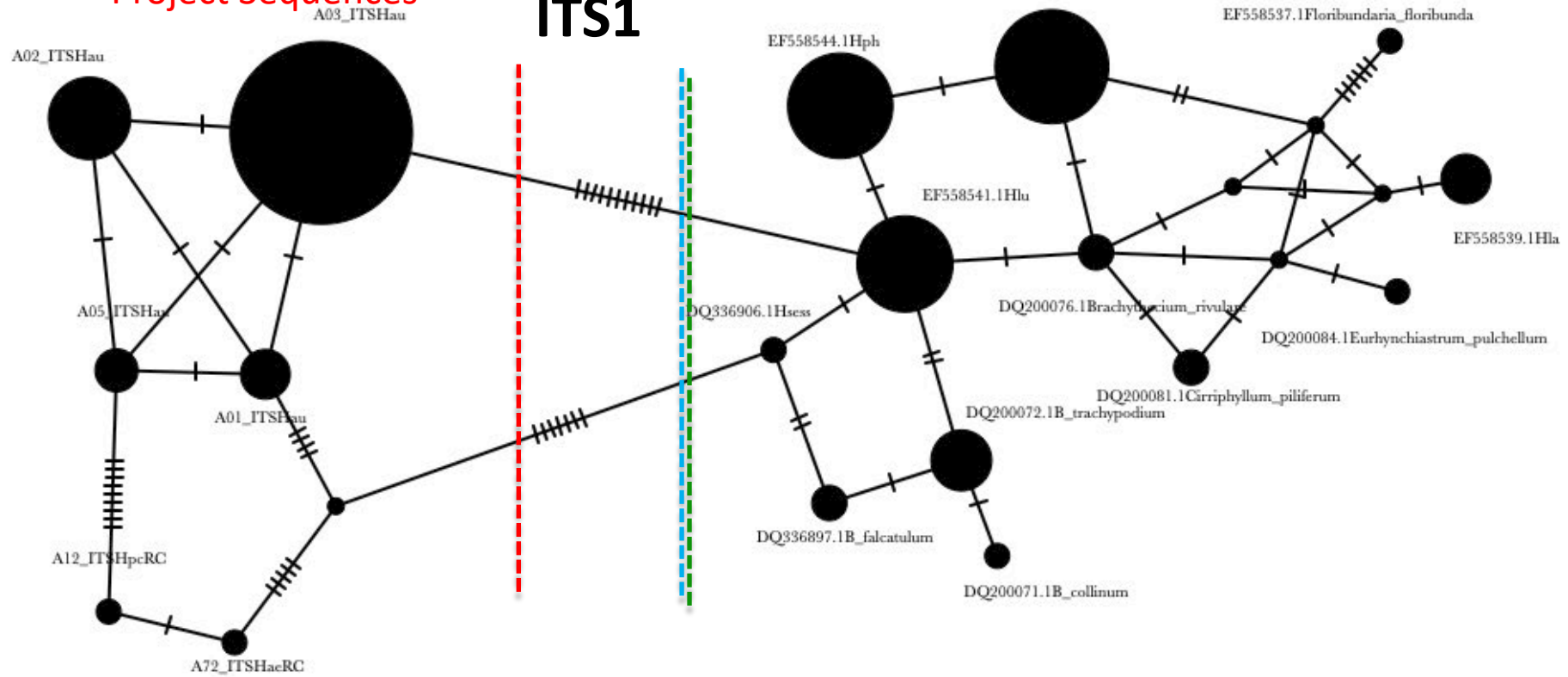
ITS2



Project Sequences

GenBank Sequences
GenBank Sequences

ITS1



Info sent by email

A. I think I understand the surprise in atpB-rbcL network and I do not believe the single nucleotide difference causing it is informative.

The atpB-rbcL grouping of a large % of my specimens with *H. nuttallii* instead of with *H. pinnafidum*/*H. aureum* is due to a single nucleotide at position 429 in the aligned atpB-rbcL. This nucleotide is sometimes a C instead of a T. If you discard all the sequences (both mine and GenBank) that are not in this basal clade of *H. nuttallii*, *H. pinnafidum*, *H. aureum* and look at the number of each of these three basal clade species that has a C instead of T, the differences between the three species are not significant. For this analysis, I am assuming that my *H. pinnafidum* specimens are all *H. pinnafidum* (so excluded the six that are clearly mis-identified).

H. pinnafidum: 31.2% (15/48) [=15 samples had C in this position instead of T, out of 48=2 GenBank + 46 my sequences]

H. aureum: 12.5% (1/8) [=1 sample had a C in this position instead of T, out of 8=2 GenBank +6 my sequences]

H. nuttallii: 0.0% (0/3) [out of 3 GenBank sequences]

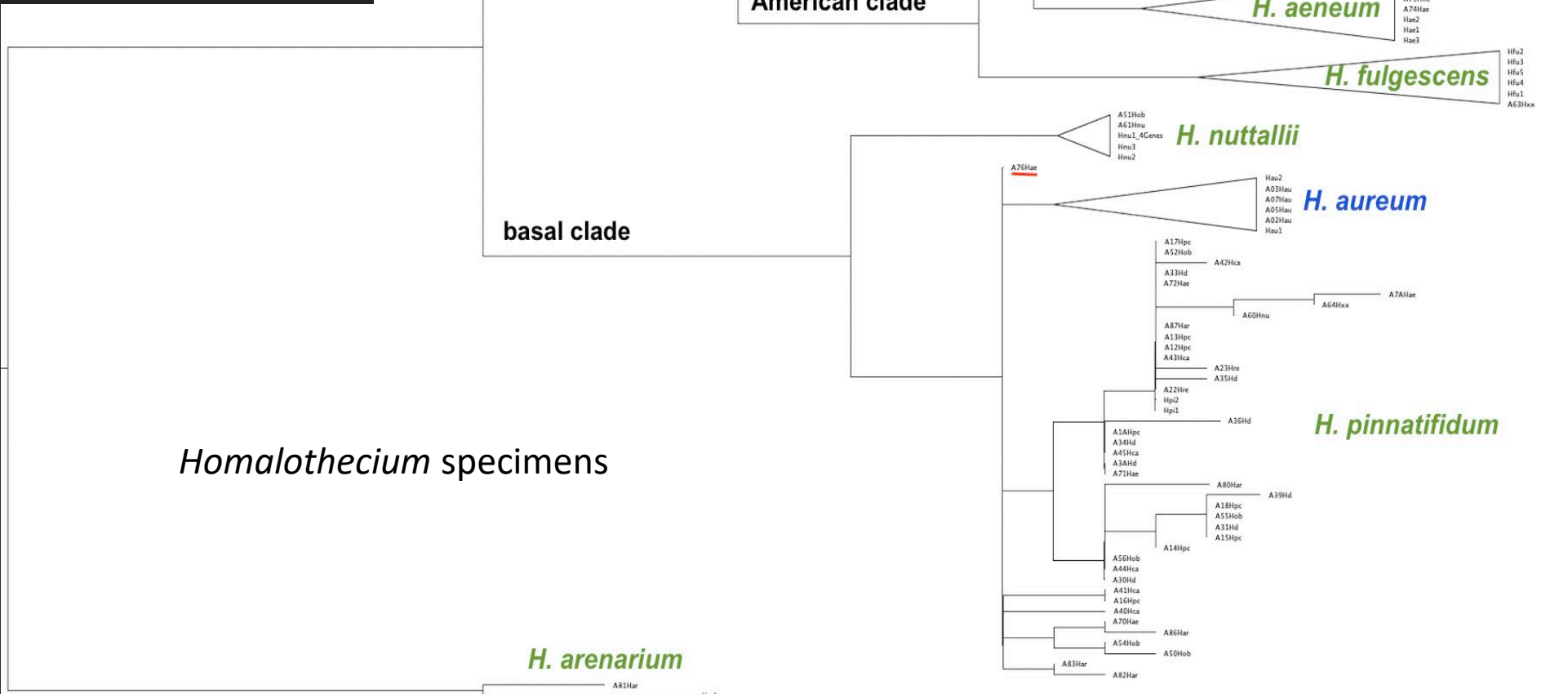
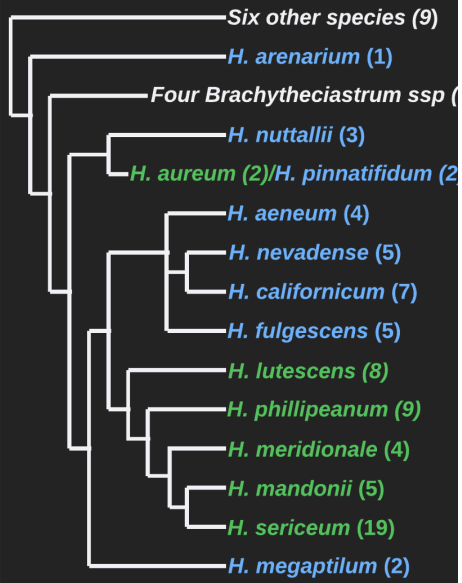
B: I think we should probably resequence both ITS1 and ITS2. If we only do one, I think it should be ITS2, since we are looking within *H. aureum*/*H. pinnafidum*. Here is what I found (see attached for details):

Number of places in gene that differentiate between	rpl16	atpB-rbcL	ITS1	ITS2
Basal clade (<i>H. aureum</i> , <i>H. pinnafidum</i> , <i>H. nuttallii</i>) from	1	1	1	7
crown clade (rest of <i>Homalothecium</i> s.str.)				
<i>H. nuttallii</i> from	0	1	4	1
<i>H. aureum</i> / <i>H. pinnafidum</i>				
<i>H. aureum</i> from	0	0	1	2
<i>H. pinnafidum</i>				

C. Since there is a leadtime, I think we should order primers and any other supplies needed now while we discuss options and decide the next best paths. The cost is relatively low and I don't mind paying for it.

D. Given the lack of/incompleteness of ITS sequences for *H. nuttallii*, if we decide to redo ITS1 and/or ITS2, I think I should find and prep a few additional *H. nuttallii* specimens for sequencing also. I don't think we need to redo the ITS sequences for the six specimens that are mis-identified. We should order now any supplies we need for that also.

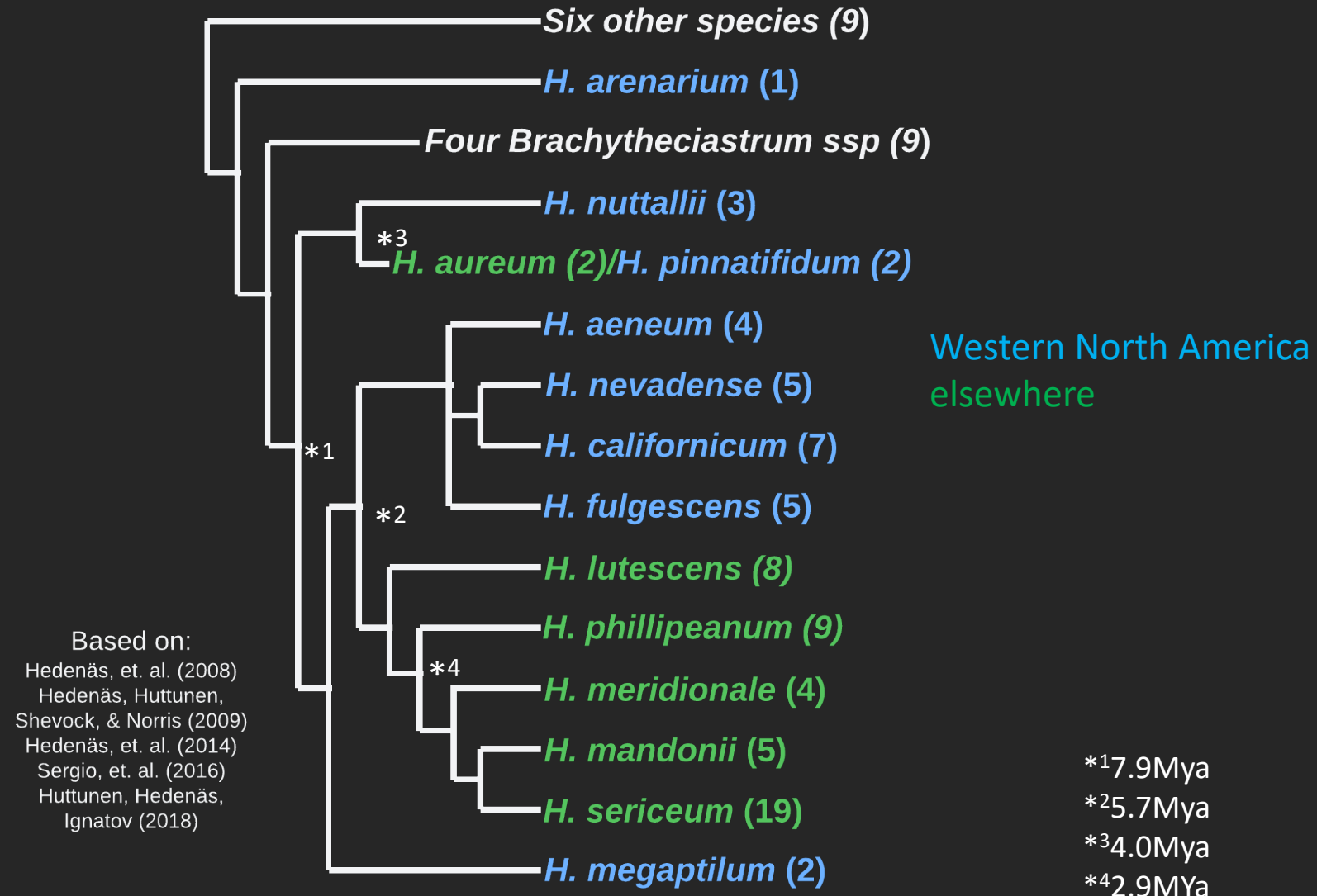
See also GeneDifferences.docx for nucleotide differences and location among these genes for GenBank sequences



230427 Meeting

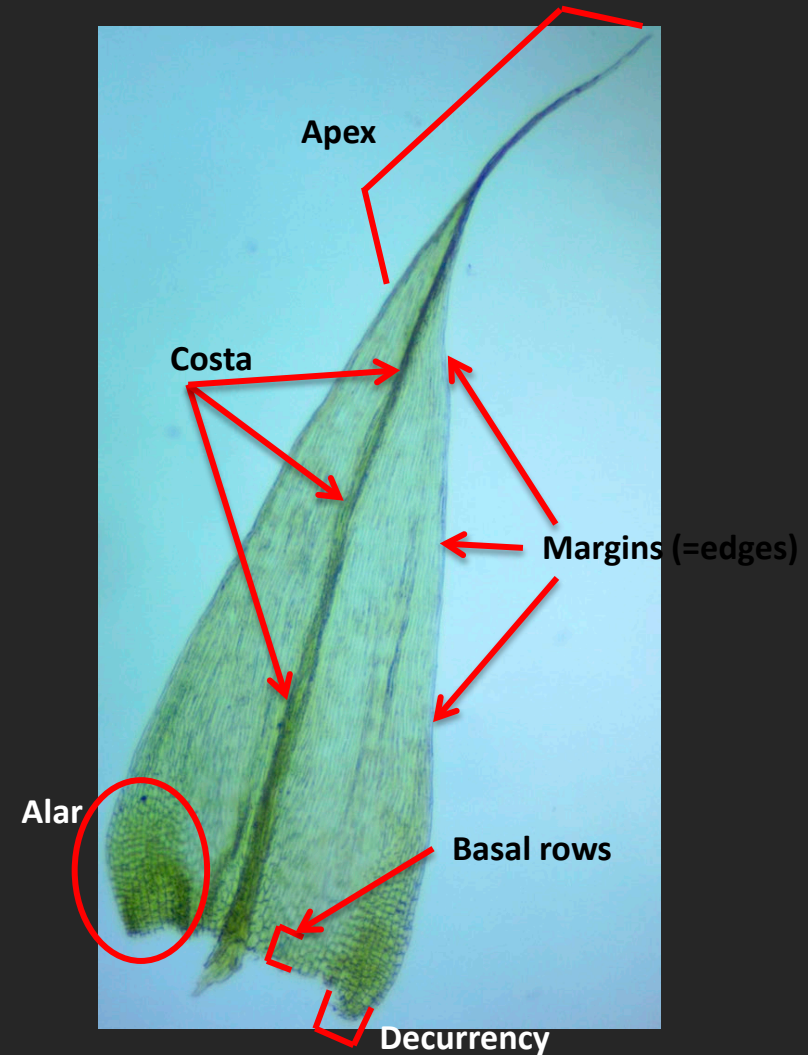
- To Do:
- Get list of specimens with both rpl16 and atpB-rbcL results in the three large nodes.
- Go through same specimens looking for morphological characters that might unite specimens in each node.

Homalothecium—Phylogeny ages



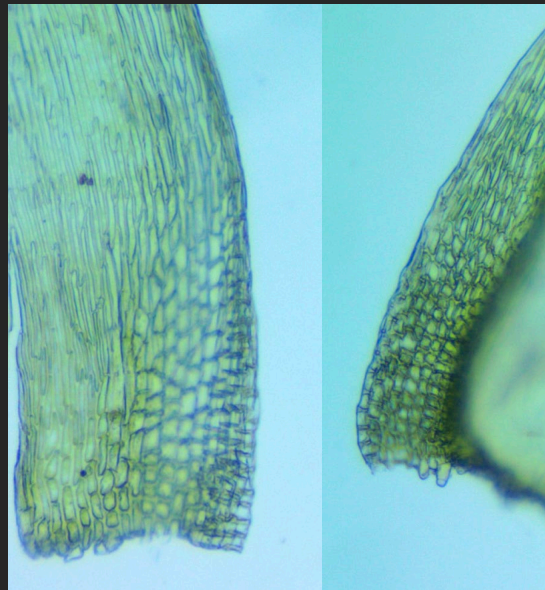
MOSS LEAF TERMINOLOGY

- Shape: overall shape of leaf
- Margin: edge of leaf
- Costa: leaf mid-rib
- Alar region: area of distinct cells in corners of leaf
- Decurrencies: cells of leaf that attach to a stem below
- Basal rows: rows of distinct cells at base of leaf
- Apex: Distal part of leaf

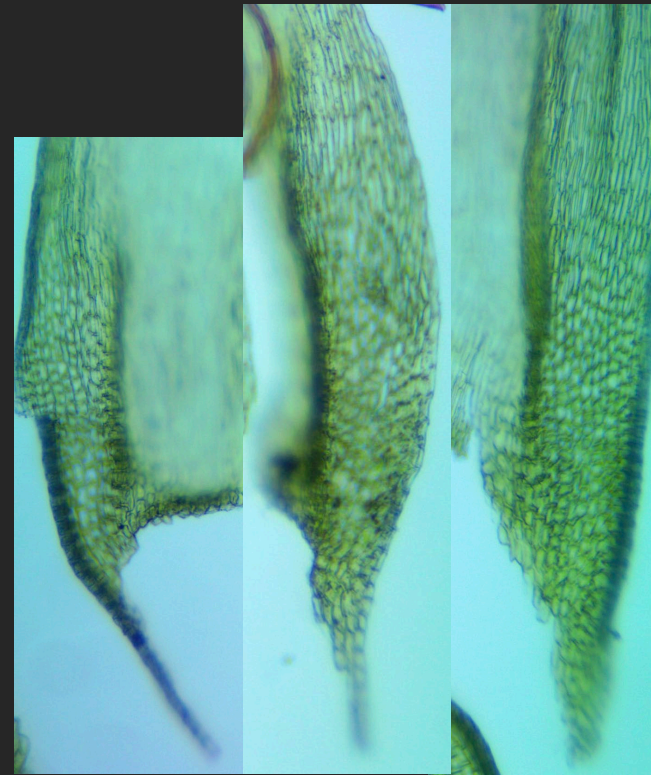


Tracked many characters

- Leaf decurrencies--# of cells (length x width)



Typical

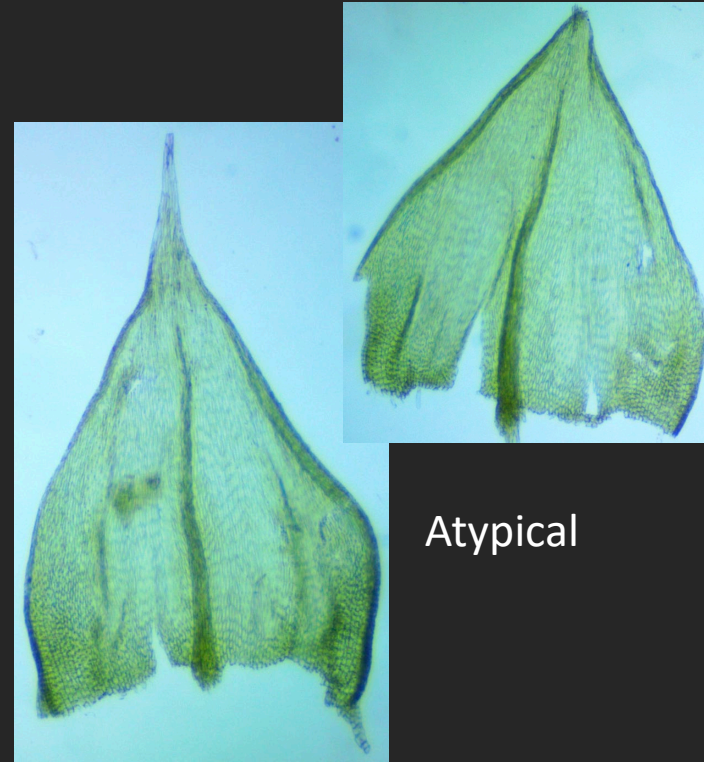


Atypical

Leaf Shape



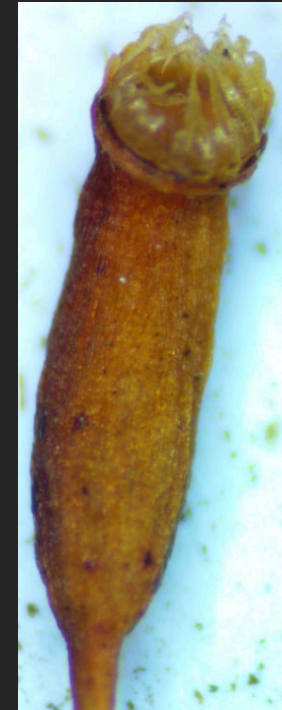
Typical



Atypical

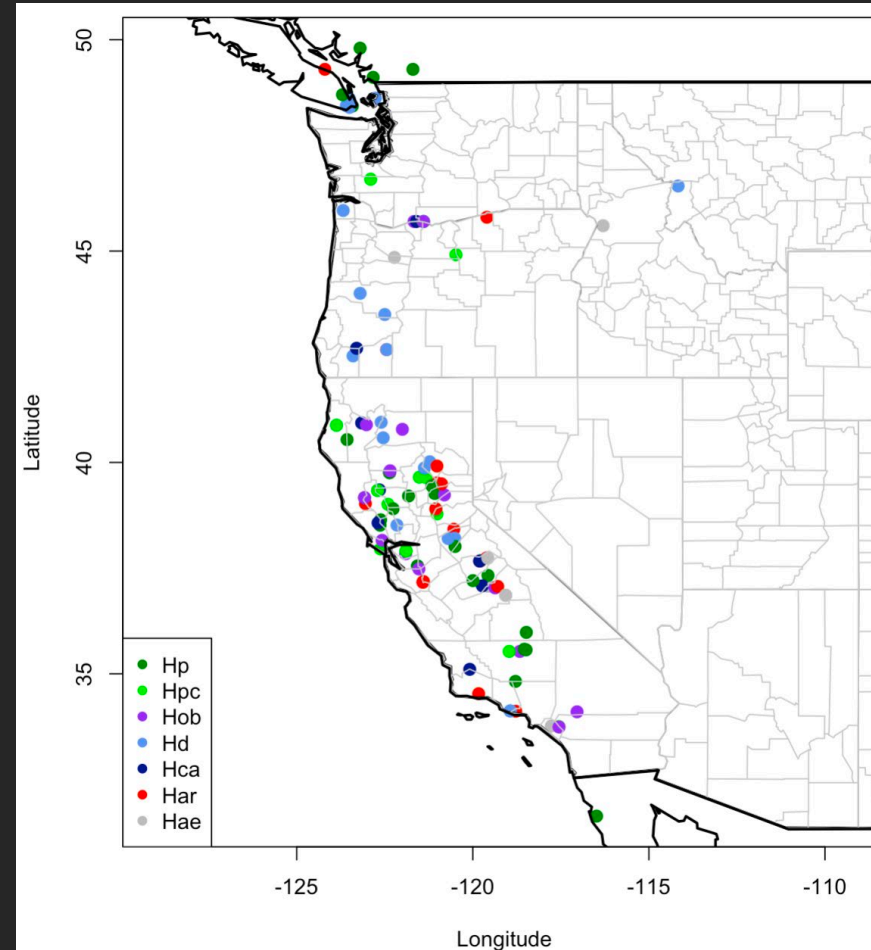
Other characters being tracked. . .

- Substrate
- Leaf length/width
- Leaf margins
- Mid-lamina cell length
- Leaf basal rows
- Leaf alar cell characteristics
- Costa spines
- Sporophyte characters
 - Operculum:
conic/rostrate/rostellate/
length
 - Urn:
amount of curvature
length
 - Seta:
rough/smooth



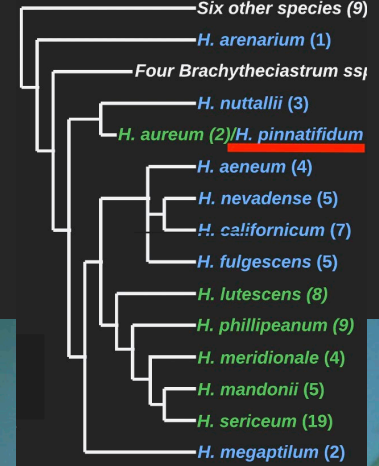
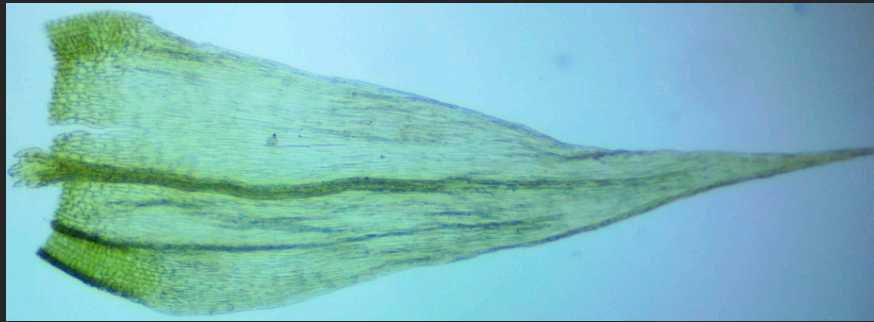
H. pinnatifidum morphotypes

- *H. pinnatifidum* specimens were classified into seven morphotypes
- Samples of each morphotype were selected for sequencing



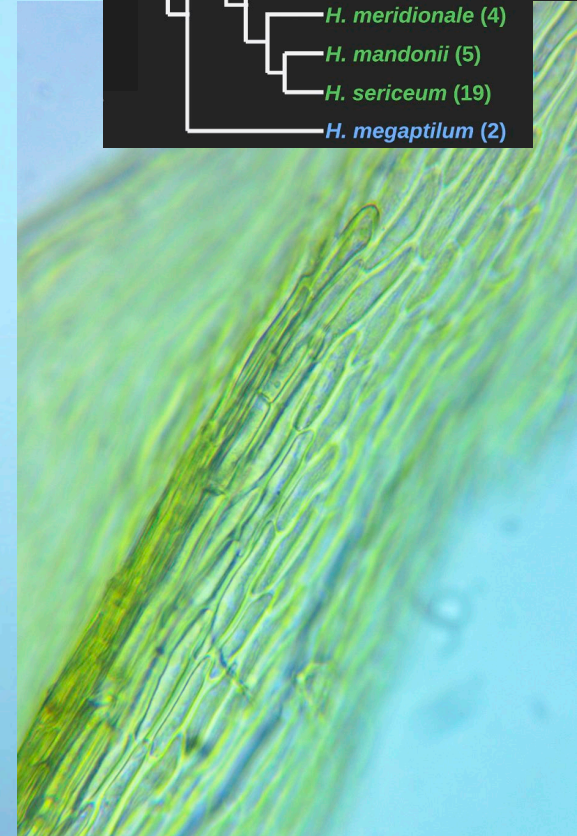
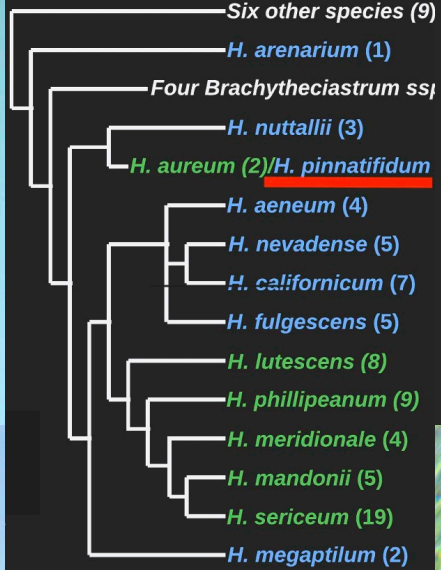
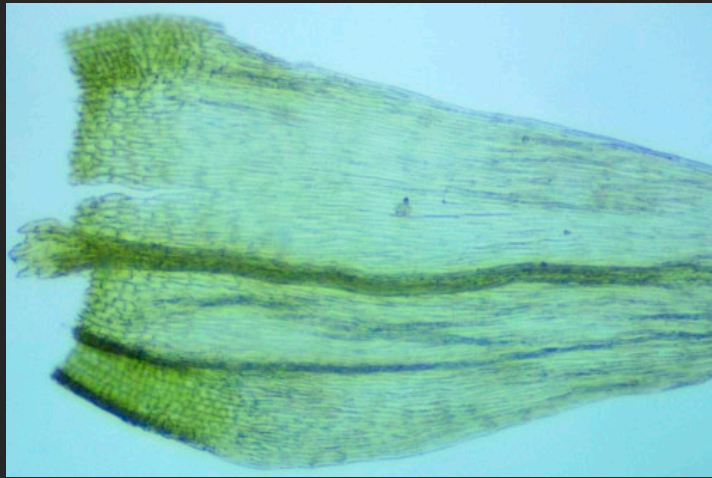
H. pinnatifidum (classic) [Hp,Hpc] A1x (10)

- Shape
- Margins
- Decurrencies
- Costa



H. pinnatifidum (classic) [Hpc, Hp]

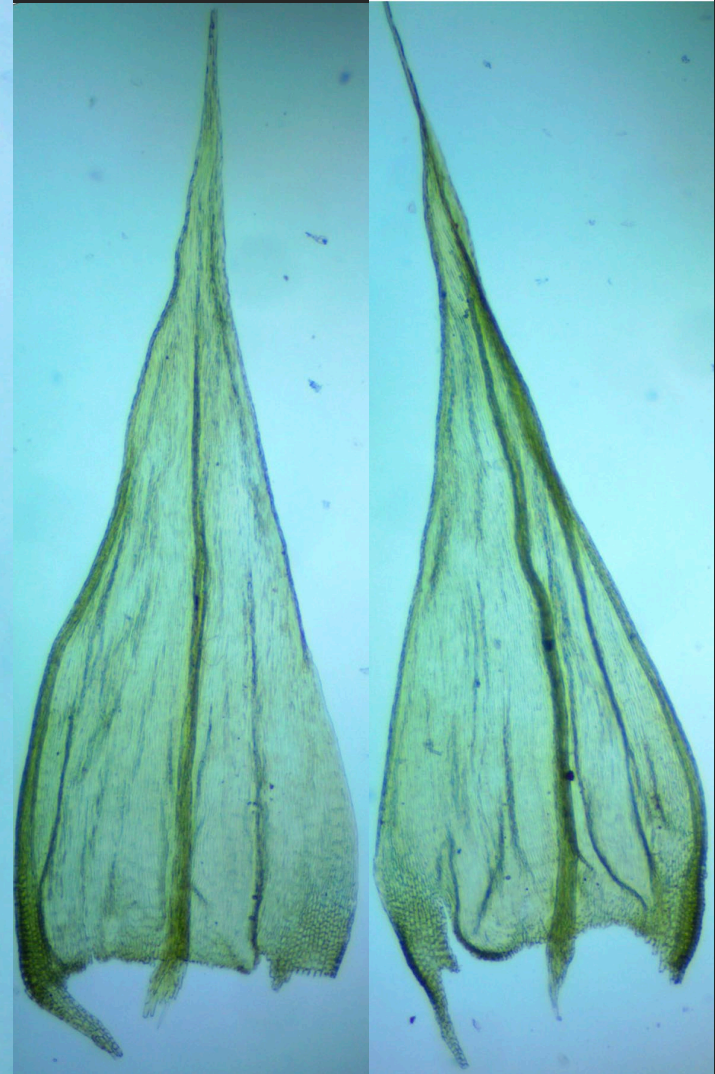
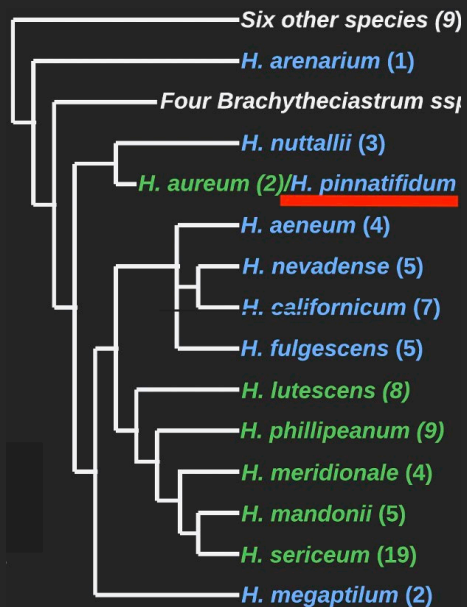
- Alar cells
- Costa spine



H. decurrentifolium

[Hd] A3x (11)

- Shape
- Costa
- Alar cells
- Margins
- Decurrencies



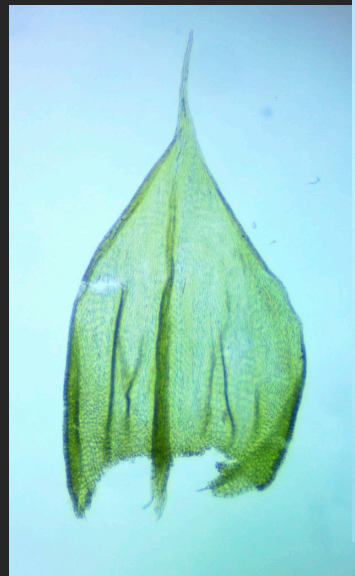
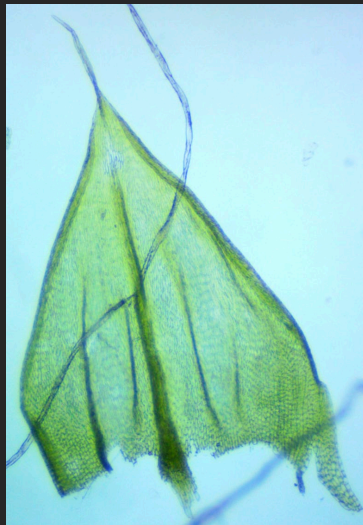
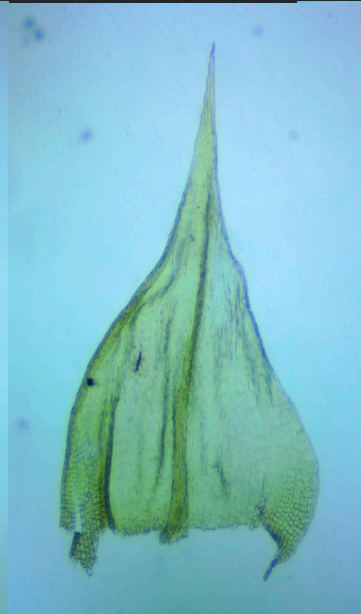
20231230.ler

H. orbicularium

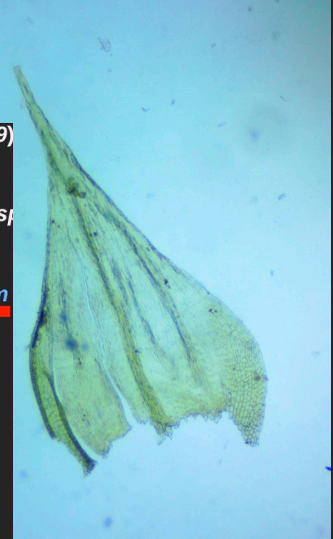
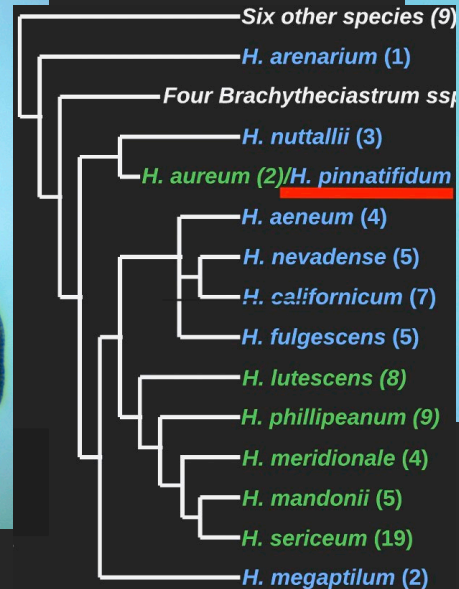
[Hob] A5x (7)



- Shape
- Costa
- Alar cells
- Margins
- Decurrencies

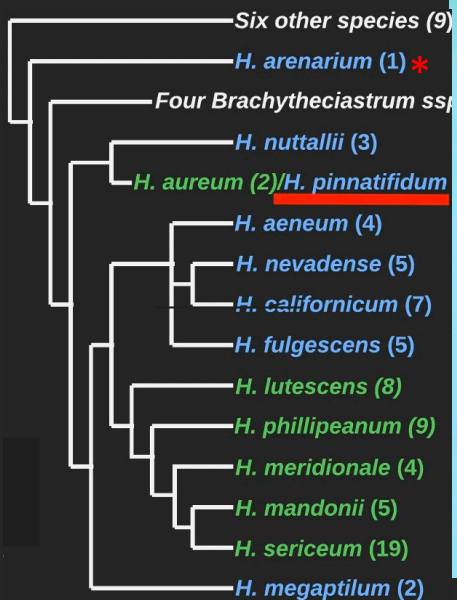
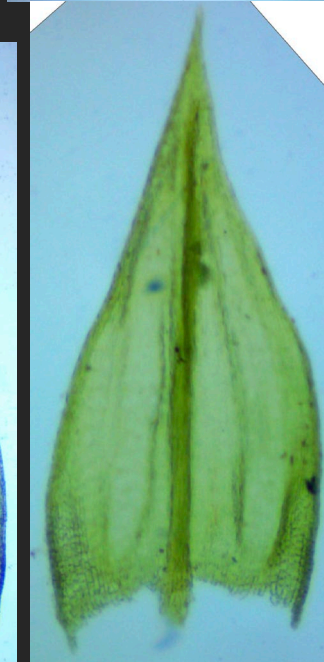
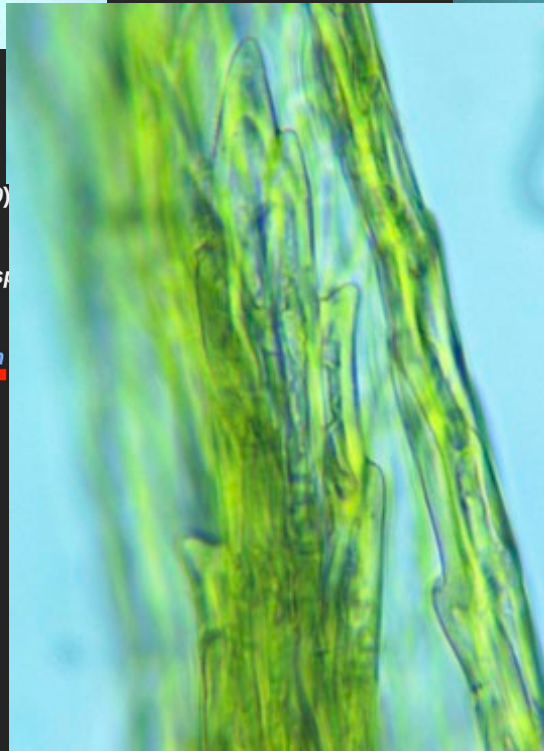


20231230.ler



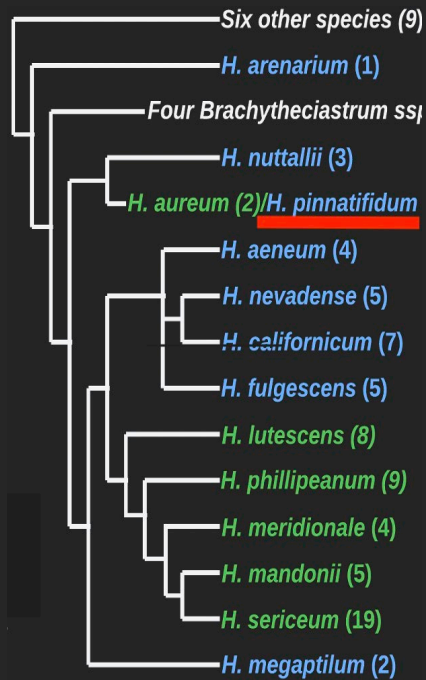
[Har] A8x (6)

- Shape
- Costa
- Alar cells
- Margins
- Decurrencies



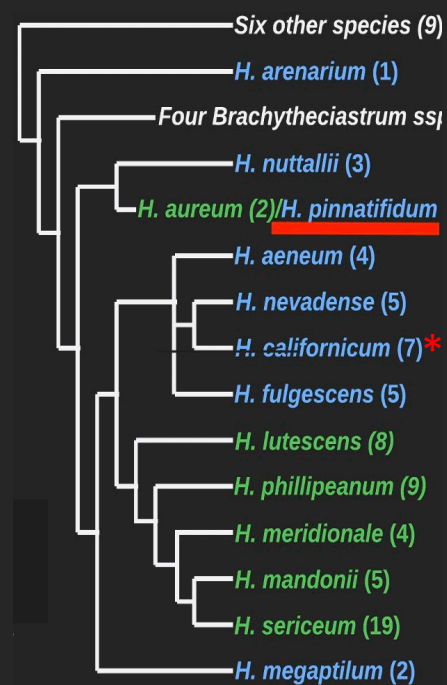
[Hre(4)]A2x

- Similar to Hpc/Hp
- Difference
 - Alar cells
 - orientation



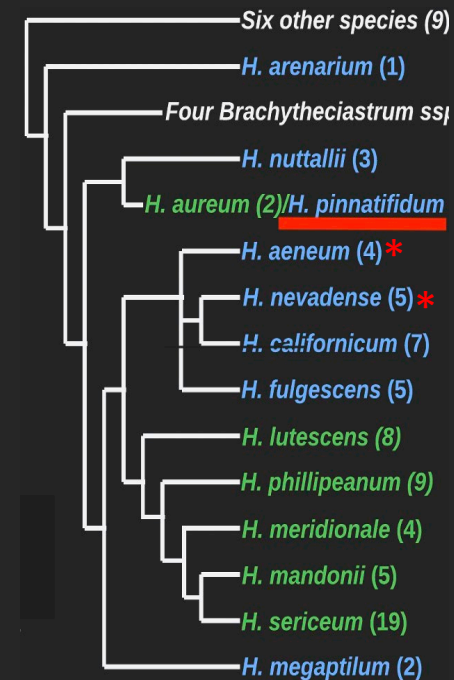
[Hca(5)]A4x

- Similar to Hd
- Differences:
 - Bigger



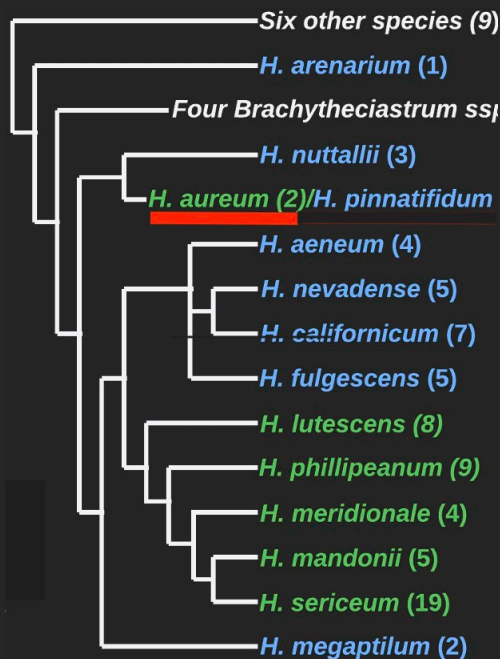
[Hae/Hne(11)]A7x

- Similar to Hpc/Hp
- Differences
 - Alar cells
 - shape/position
 - number
 - walls



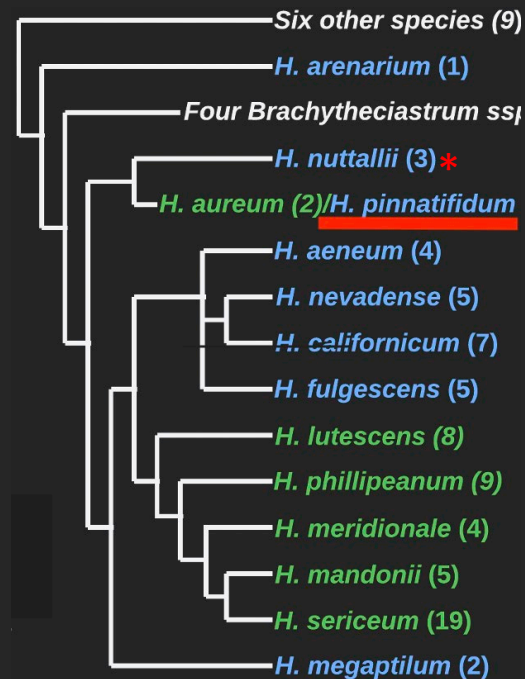
[Hau(7)]

- Not North America
 - North Africa(2)
 - Southern Europe(4)
 - East Kazakhstan(1)
- (4)



[Hnu(3) / Hxx(2)]

- 1 IDed as *H. nuttallii*
- 2 with a few very minute recurved margin teeth
- ID=*H. pinnatifidum*
- Did not fit another morphotype



H. pinnatifidum morphotypes leaf length/width ratio

