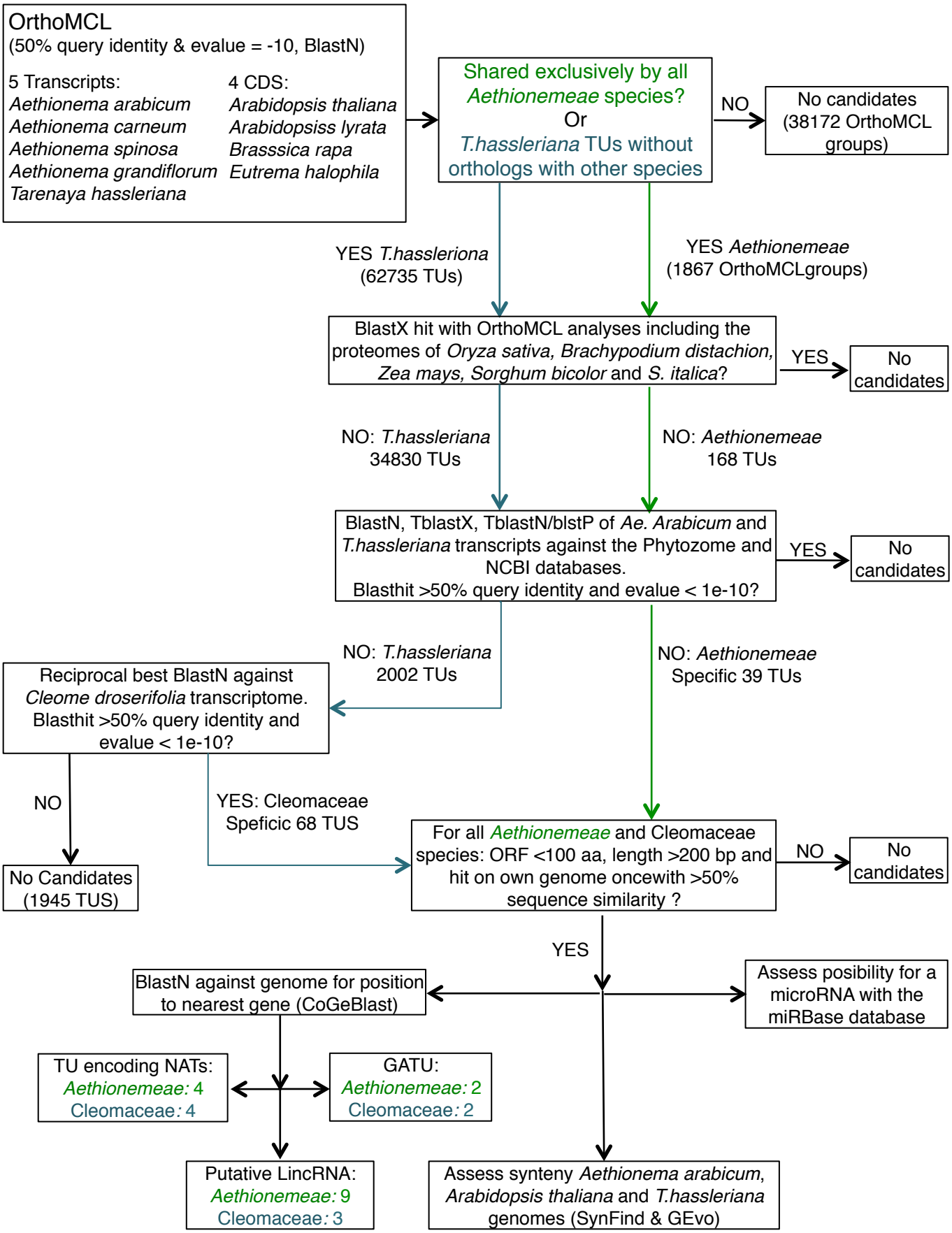
OrthoMCL
(50% query identity & evalue = -10, BlastN)
5 Transcripts:
Aethionema arabicum
Aethionema carneum
Aethionema spinosa
Aethionema grandiflorum
Tarenaya hassleriana
4 CDS:
Arabidopsis thaliana
Arabidopsiss lyrata
Brasssica rapa
Eutrema halophila
Shared exclusively by all Aethionemeae species?
Or
T.hassleriana TUs without orthologs with other species
NO
No candidates (38172 OrthoMCL groups)
YES T.hassleriona (62735 TUs)
YES Aethionemeae (1867 OrthoMCLgroups)
BlastX hit with OrthoMCL analyses including the proteomes of Oryza sativa, Brachypodium distachion, Zea mays, Sorghum bicolor and S. italica?
YES
No candidates
NO: T.hassleriana 34830 TUs
NO: Aethionemeae 168 TUs
BlastN, TblastX, TblastN/blstP of Ae. Arabicum and T.hassleriana transcripts against the Phytozome and NCBI databases.
Blasthit >50% query identity and evalue < 1e-10?
YES
No candidates
NO: T.hassleriana 2002 TUs
NO: Aethionemeae Specific 39 TUs
Reciprocal best BlastN against Cleome droserifolia transcriptome.
Blasthit >50% query identity and evalue < 1e-10?
NO
No Candidates (1945 TUS)
YES: Cleomaceae Speficic 68 TUS
For all Aethionemeae and Cleomaceae species: ORF <100 aa, length >200 bp and hit on own genome oncewith >50% sequence similarity ?
NO
No candidates
YES
BlastN against genome for position to nearest gene (CoGeBlast)
Assess posibility for a microRNA with the miRBase database
TU encoding NATs:
Aethionemeae: 4
Cleomaceae: 4
GATU:
Aethionemeae: 2
Cleomaceae: 2
Putative LincRNA:
Aethionemeae: 9
Cleomaceae: 3
Assess synteny Aethionema arabicum, Arabidopsis thaliana and T.hassleriana genomes (SynFind & GEvo)