doi:10.1038/nature12130

Supplementary Table 1. The Taxonomy of the Organisms Used in this Study

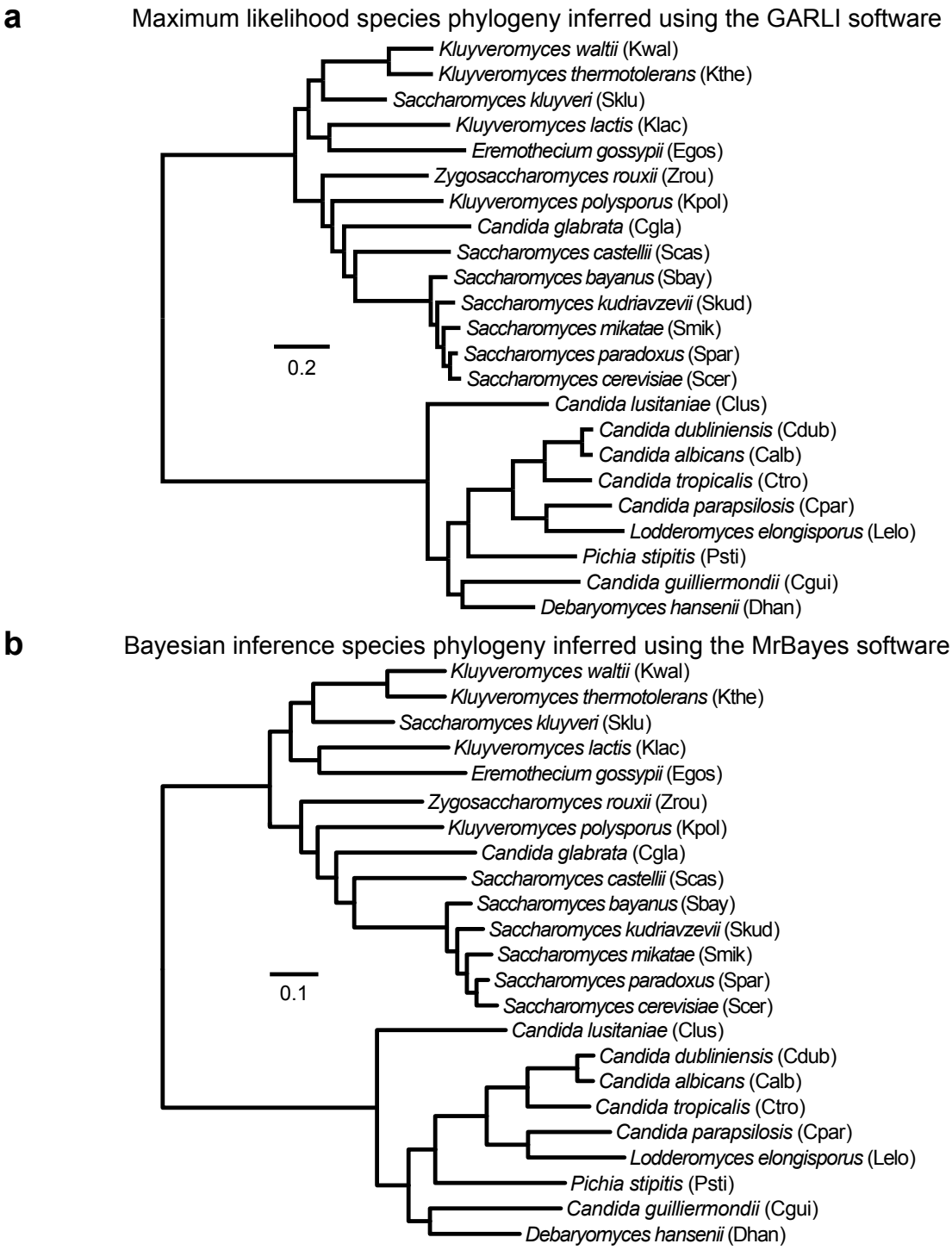
Organism (acronym)	Taxonomy
Yeasts	
Kluyveromyces waltii (Kwal)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Kluyveromyces thermotolerans (Kthe)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Saccharomyces kluyveri (Sklu)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Kluyveromyces lactis (Klac)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Eremothecium gossypii (Egos)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Zygosacharomyces rouxii (Zrou)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Kluyveromyces polysporus (Kpol)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Candida glabrata (Cgla)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Saccharomyces castellii (Scas)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Saccharomyces bayanus (Sbay)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Saccharomyces kudriavzevii (Skud)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Saccharomyces mikatae (Smik)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Saccharomyces paradoxus (Spar)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Saccharomyces cerevisiae (Scer)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Candida lusitaniae (Clus)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Candida dubliniensis (Cdub)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Candida albicans (Calb)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Candida tropicalis (Ctro)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Candida parapsilosis (Cpar)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Lodderomyces elongisporus (Lelo)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae;
Pichia stipitis (Psti)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Candida guilliermondii (Cgui)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Debaryomyces hansenii (Dhan)	Fungi;Ascomycota;Saccharomycetes;Saccharomycetaceae
Verterbrates	
Xenopus tropicalis (Xtro)	Animalia;Chordata;Aphibia;Anura;Pipidae
Gallus gallus (Ggal)	Animalia;Chordata;Aves;Galliformes;Phasianidae
Monodelphis domestica (Mdom)	Animalia;Chordata;Monodelphis;Mammalia;Didelphimorphia
Bos taurus (Btau)	Animalia;Chordata;Mammalia;Artiodactyla;Bovidae
Equus caballus (Ecab)	Animalia;Chordata;Mammalia;Perissodactyla;Equidae
Canis familiaris (Cfam)	Animalia;Chordata;Mammalia;Carnivora;Canidae
Macaca mulatta (Mmul)	Animalia;Chordata;Mammalia;Primates;Cercopithecidae
Pongo pygmaeus (Ppyg)	Animalia;Chordata;Mammalia;Primates;Hominidae
Homo sapiens (Hsap)	Animalia;Chordata;Mammalia;Primates;Hominidae
*Pan troglodytes (*Ptro)	Animalia;Chordata;Mammalia;Primates;Hominidae
Rattus norvegicus (Rnor)	Animalia;Chordata;Mammalia;Rodentia;Muridae
Mus musculus (Mmus)	Animalia;Chordata;Mammalia;Rodentia;Muridae
Cavia porcellus (Cpor)	Animalia;Chordata;Mammalia;Rodentia;Caviidae
Danio rerio (Drer)	Animalia;Chordata;Actinopterygii;Cypriniformes;Cyprinidae

Oryzias latipes (Olat)	Animalia;Chordata;Actinopterygii;Beloniformes;Adrianichthyidae
Tetraodon nigroviridis (Tnig)	Animalia;Chordata;Actinopterygii;Tetraodontiformes;Tetraodontidae
Takifugu rubripes (Trub)	Animalia;Chordata;Actinopterygii;Tetraodontiformes;Tetraodontidae
Gasterosteus aculeatus (Gacu)	Animalia;Chordata;Actinopterygii;Gasterosteiformes;Gasterosteidae
Metazoa	
Strongylocentrotus purpuratus (Spur)	Animalia;Echinodermata;Echinoidea;Echinoida;Strongylocentrotidae
Branchiostoma floridae (Bflo)	Animalia;Chordata;Leptocardii;Amphioxiformes;Branchiostomidae
Ciona intestinalis (Cint)	Animalia;Chordata;Ascidiaceae;Enterogona;Cionidae
Mus musculus (Mmus)	Animalia;Chordata;Mammalia;Rodentia;Muridae
Gallus gallus (Ggal)	Animalia;Chordata;Aves;Galliformes;Phasianidae
Homo sapiens (Hsap)	Animalia;Chordata;Mammalia;Primates;Hominidae
Xenopus tropicalis (Xtro)	Animalia;Chordata;Aphibia;Anura;Pipidae
Danio rerio (Drer)	Animalia;Chordata;Actinopterygii;Cypriniformes;Cyprinidae
Helobdella robusta (Hrob)	Animalia;Annelida;Clitellata;Rhynchobdellida;Glossiphoniidae
Lottia gigantea (Lgig)	Animalia;Mollusca;Gastropoda;Patellogastropoda;Lottiidae
Caenorhabditis elegans (Cele)	Animalia;Nematoda;Secernentea;Rhabditida;Rhabditidae
Schistosoma mansoni (Sman)	Animalia;Platyhelminthes;Digenea;Strigeidida
Ixodes scapularis (Isca)	Animalia;Arthropoda;Arachnida;Ixodida;Ixodidae
Daphnia pulex (Dpul)	Animalia;Arthropoda;Branchiopoda;Cladocera;Daphniidae
Apis mellifera (Amel)	Animalia;Arthropoda;Insecta;Hymenoptera;Apidae
Tribolium castaneum (Tcas)	Animalia;Arthropoda;Insecta;Coleoptera;Tenebrionidae
Drosophila melanogaster (Dmel)	Animalia;Arthropoda;Insecta;Diptera;Drosophilidae
Bombyx mori (Bmor)	Animalia;Arthropoda;Insecta;Lepidopteroa;Bombycidae
Monosiga brevicollis (Mbre)	hoanoflagellida;Codonosigidae
Nematostella vectensis (Nvec)	Animalia;Cnidaria;Anthozoa;Actiniaria;Edwardsiidae
Trichoplax adhaerens (Tadh)	Animalia;Placozoa;Tricoplacia;Tricoplaciformes;Trichoplacidae

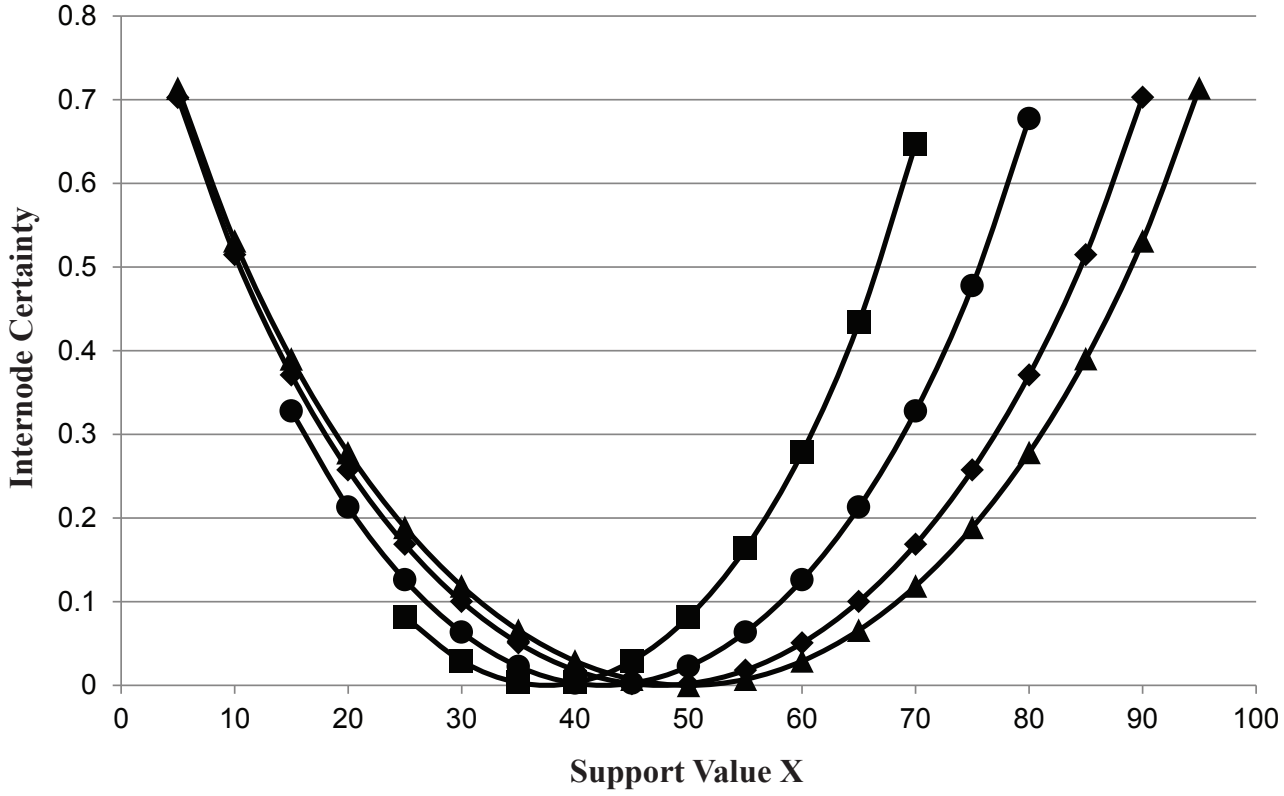
Supplementary Table 2.Bipartitions that Significantly Conflict with the Bipartitions Recovered in the Concatenation Phylogeny of 23 Yeast Genomes

Primary Tree Bipartition	GSF	IC	[Conflicting Bipartition]:GSF value
Kthe, Kwal	99	0.96	None
Calb, Cdub	98	0.95	None
Sbay, Scer, Skud, Smik, Spar	99	0.97	None
Calb, Cdub, Cgui, Clus, Cpar, Ctro, Dhan, Lelo, Psti	95	0.90	None
Calb, Cdub, Ctro	90	0.78	None
Cpar, Lelo	89	0.77	None
Calb, Cdub, Cpar, Ctro, Lelo	86	0.76	None
Scer, Spar	77	0.54	[Sbay,Skud,Smik,Spar]:8; [Smik,Spar]:5
Cgla, Kpol, Sbay, Scas, Scer, Skud, Smik, Spar, Zrou	62	0.59	[Calb,Cdub,Cgla,Cgui,Clus,Cpar,Ctro,Dhan,Lelo,Psti]:6; [Calb,Cdub,Cgui,Clus,Cpar,Ctro,Dhan,Lelo,Psti,Zrou]:5
Scer, Smik, Spar	60	0.30	[Sbay,Skud,Smik]:14; [Sbay,Scer,Skud,Spar]:11; [Sbay,Skud,Smik,Spar]:8; [Skud,Smik]:7 ; [Scer,Skud,Spar]:6
Scer, Skud, Smik ,Spar	52	0.06	[Sbay,Skud]:29; [Sbay,Skud,Smik]:14; [Sbay,Scer,Smik,Spar]:11; [Sbay,Scer,Skud,Spar]:11; [Sbay,Skud,Smik,Spar]:8
Calb, Cdub, Cpar, Ctro, Lelo, Psti	45	0.11	[Cgui,Clus,Dhan,Psti]:20; [Dhan,Psti]:11; [Cgui,Dhan,Psti]:10; [Calb,Cdub,Cgui,Clus,Cpar,Ctro,Dhan,Lelo]:8; [Calb,Cdub,Clus,Cpar,Ctro,Lelo]:5; [Clus,Dhan,Psti]:5
Kthe, Kwal, Sklu	41	0.32	[Agos,Kthe,Kwal]:9; [Calb,Cdub,Cgui,Clus,Cpar,Ctro,Dhan,Kthe, Kwal,Lelo,Psti]:9; [Klac,Sklu]:8; [Agos,Klac,Kthe,Kwal]:7; [Agos, Klac,Sklu]:7; [Agos,Sklu]:7; [Cgla,Kpol,Kthe,Kwal,Sbay,Scas, Scer,Skud,Smik,Spar,Zrou]:6; [Klac,Kthe,Kwal]:5; [Cgla,Kpol, Sbay,Scas,Scer,Sklu,Skud,Smik,Spar,Zrou]:5
Agos, Klac	36	0.09	[Agos,Cgla,Kpol,Kthe,Kwal,Sbay,Scas,Scer,Sklu,Skud,Smik,Spar, Zrou]:17; [Cgla,Klac,Kpol,Kthe,Kwal,Sbay,Scas,Scer,Sklu,Skud, Smik,Spar,Zrou]:13; [Agos,Kthe,Kwal,Sklu]:13; [Klac,Kthe,Kwal, Sklu]:10; [Agos,Kthe,Kwal]:9; [Klac,Sklu]:8; [Agos,Sklu]:7; [Cgla, Klac,Kpol,Sbay,Scas,Scer,Skud,Smik,Spar,Zrou]:7; [Klac,Kthe, Kwal]:5
Agos, Klac, Kthe, Kwal, Sklu	31	0.04	[Agos,Calb,Cdub,Cgui,Clus,Cpar,Ctro,Dhan,Klac,Lelo,Psti]:19; [Calb,Cdub,Cgui,Clus,Cpar,Ctro,Dhan,Klac,Lelo,Psti]:17; [Agos, Calb,Cdub,Cgui,Clus,Cpar,Ctro,Dhan,Lelo,Psti]:13; [Calb,Cdub, Cgui,Clus,Cpar,Ctro,Dhan,Kthe,Kwal,Lelo,Psti]:9; [Agos,Cgla, Klac,Kpol,Sbay,Scas,Scer,Skud,Smik,Spar,Zrou]:7; [Cgla,Klac, Kpol,Sbay,Scas,Scer,Skud,Smik,Spar,Zrou]:7; [Cgla,Kpol,Kthe, Kwal,Sbay,Scas,Scer,Skud,Smik,Spar,Zrou]:6; [Cgla,Kpol,Sbay, Scas,Scer,Sklu,Skud,Smik,Spar,Zrou]:5
Cgla, Sbay, Scas, Scer, Skud, Smik, Spar	29	0.12	[Cgla,Kpol]:12; [Kpol,Scas]:10; [Kpol,Sbay,Scas,Scer,Skud,Smik, Spar,Zrou]:9; [Kpol,Sbay,Scas,Scer,Skud,Smik,Spar]:8; [Cgla,

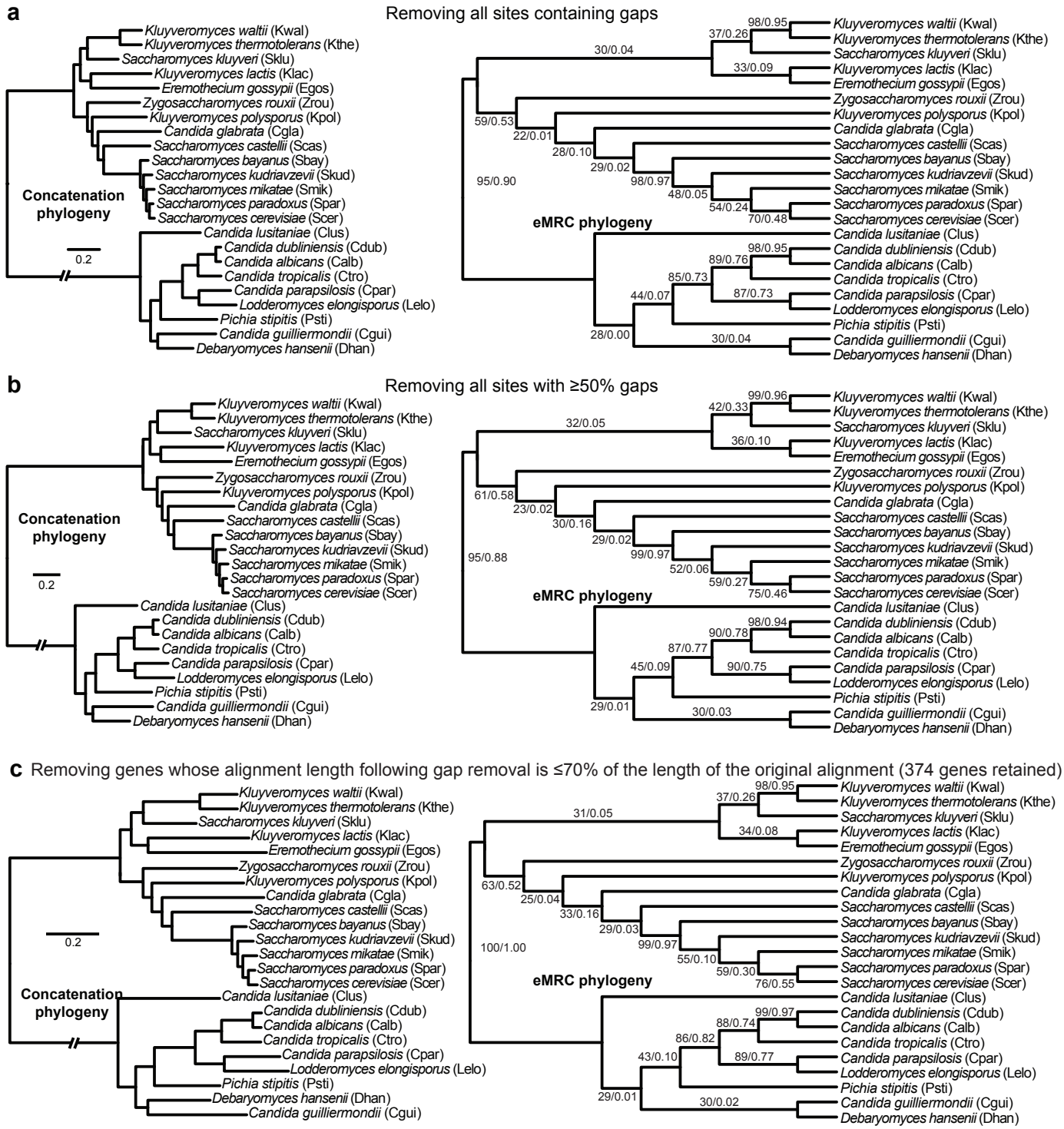
			Zrou]:8; [Kpol,Sbay,Scer,Skud,Smik,Spar]:8; [Sbay,Scer,Skud, Smik,Spar,Zrou]:7; [Sbay,Scas,Scer,Skud,Smik,Spar,Zrou]:7; [Cgla,Kpol,Scas]:6; [Agos,Klac,Kpol,Kthe,Kwal,Sbay,Scas,Scer, Sklu,Skud,Smik,Spar,Zrou]:6; [Scas,Zrou]:5
Sbay, Scas, Scer, Skud, Smik, Spar	29	0.02	[Cgla,Sbay,Scer,Skud,Smik,Spar]:20; [Cgla,Scas]:17; [Kpol, Scas]:10; [Kpol,Sbay,Scer,Skud,Smik,Spar]:8; [Sbay,Scer,Skud, Smik,Spar,Zrou]:7; [Cgla,Kpol,Scas]:6; [Scas,Zrou]:5
Calb, Cdub, Cgui, Cpar, Ctro, Dhan, Lelo, Psti	29	0.01	[Cgui,Clus,Dhan]:24; [Cgui,Clus,Dhan,Psti]:20; [Cgui,Clus]:20; [Calb,Cdub,Clus,Cpar,Ctro,Dhan,Lelo,Psti]:16; [Clus,Dhan]:12; [Calb,Cdub,Clus,Cpar,Ctro,Lelo,Psti]:9; [Calb,Cdub,Cgui,Clus, Cpar,Ctro,Dhan,Lelo]:8; [Calb,Cdub,Cgui,Clus,Cpar,Ctro,Lelo, Psti]:6; [Clus,Dhan,Psti]:5; [Calb,Cdub,Clus,Cpar,Ctro,Lelo]:5
Cgui, Dhan	29	0.02	[Cgui,Clus]:20; [Calb,Cdub,Cpar,Ctro,Dhan,Lelo,Psti]:18; [Calb,Cdub,Clus,Cpar,Ctro,Dhan,Lelo,Psti]:16; [Clus,Dhan]:12; [Dhan,Psti]:11; [Calb,Cdub,Cgui,Cpar,Ctro,Lelo,Psti]:6; [Calb, Cdub,Cgui,Clus,Cpar,Ctro,Lelo,Psti]:6; [Clus,Dhan,Psti]:5
Cgla, Kpol, Sbay, Scas, Scer, Skud, Smik, Spar	24	0.02	[Kpol,Zrou]:17; [Cgla,Sbay,Scas,Scer,Skud,Smik,Spar,Zrou]:15; [Kpol,Sbay,Scas,Scer,Skud,Smik,Spar,Zrou]:9; [Cgla,Zrou]:8; [Sbay,Scer,Skud,Smik,Spar,Zrou]:7; [Sbay,Scas,Scer,Skud,Smik, Spar,Zrou]:7; [Calb,Cdub,Cgla,Cgui,Clus,Cpar,Ctro,Dhan,Lelo, Psti]:6; [Scas,Zrou]:5



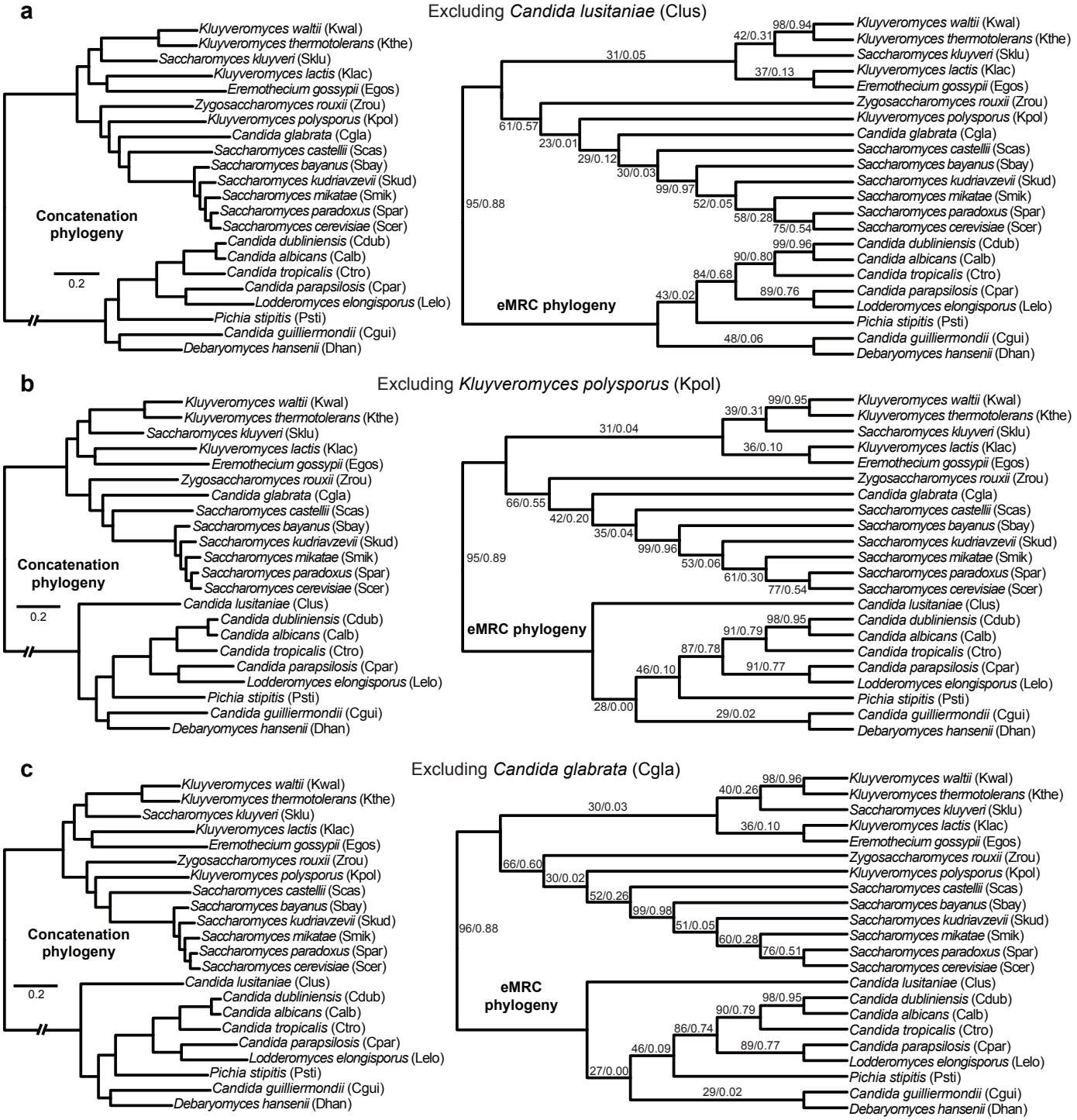
Supplementary Figure 1 | The topology of the yeast phylogeny recovered from concatenation analyses using one other maximum likelihood software (GARLI) and one other Bayesian inference (MrBayes) software was identical to the topology recovered by maximum likelihood analysis using the RAxML software. a, The yeast species phylogeny recovered from concatenation analysis of 1,070 genes using maximum likelihood as implemented in the GARLI software. All internodes received 100% bootstrap support. **b**, The yeast species phylogeny recovered from concatenation analysis of 1,070 genes using Bayesian inference as implemented in the MrBayes software. All internodes had 100% posterior probability.



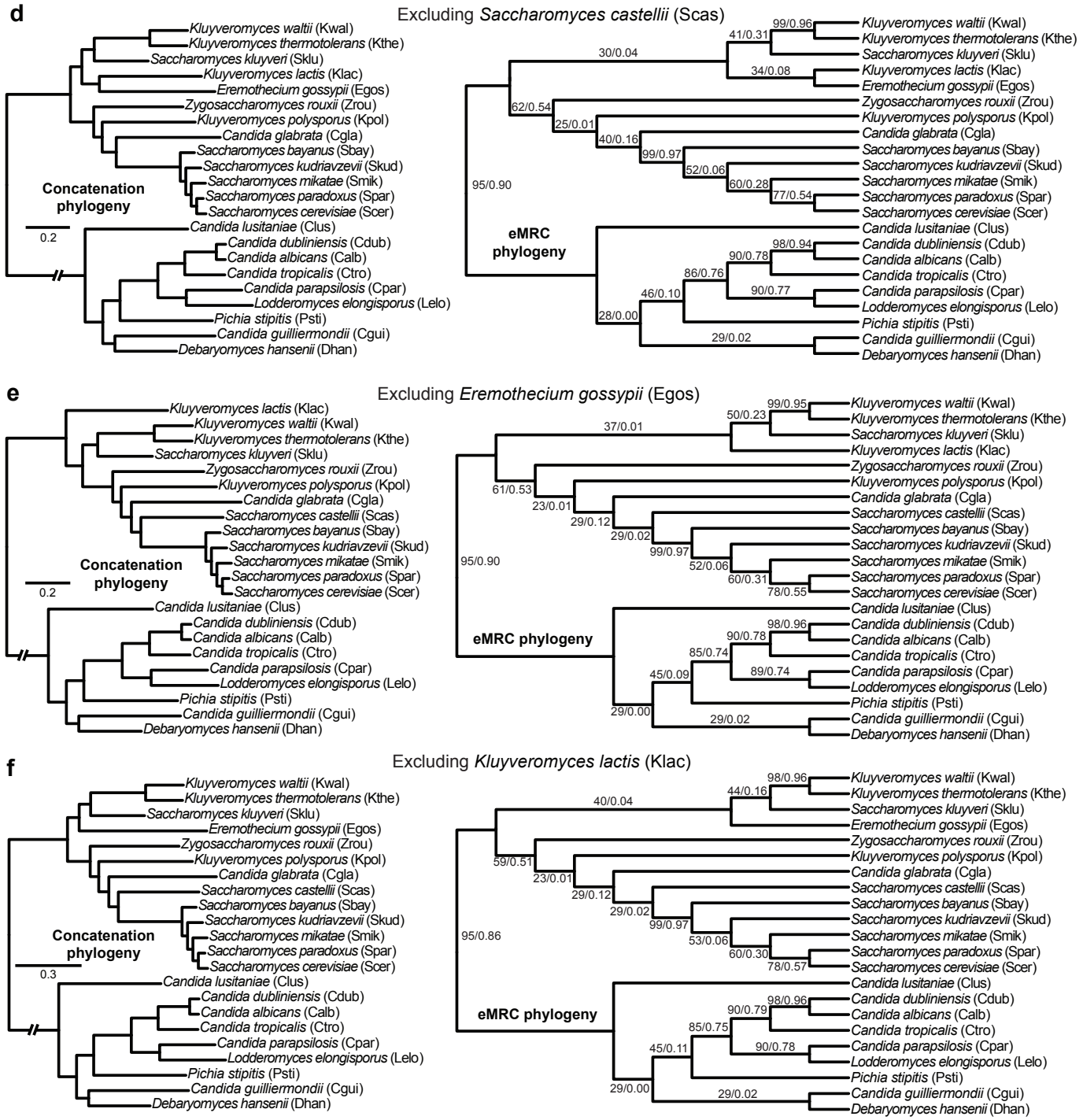
Supplementary Figure 2 | Representative values of the new measure Internode Certainty (IC) for a range of representative support values of two most prevalent and conflicting bipartitions for a given internode. Each plot on the graph depicts how IC (Y-axis) varies in response to the relative support of conflicting bipartitions on a given internode. IC can be measured on any given set of trees. For example, if the entire set of gene trees (GTs) is used, the IC value of a given internode will reflect the amount of information available for that internode in the set GTs by considering the internode's gene support frequency jointly with the frequency of the most prevalent bipartition that conflicts with the internode. If the set of bootstrap replicate trees for a given gene is used, then IC will be calculated based on bootstrap support values. From the right to the left of the graph, the first of the four plots shown with triangle symbols corresponds to the case of only two conflicting bipartitions for one internode with support values X and 100–X. For example, given 100 total GTs, if 60 of them support bipartition 1, the remaining 40 will support the conflicting bipartition. The second, third and fourth of the four plots (shown with diamond, circle, and square symbols, respectively) correspond to case where there are three conflicting bipartitions for one internode, but only the two most prevalent ones are considered. For example, in the plot with the diamond symbols, given 100 total GTs, if 60 of them support bipartition 1, 35 will support the conflicting bipartition 2, because conflicting bipartition 3 has been set to be supported by 5 GTs. Thus, when the two most prevalent ones are considered, the percentage of GTs supporting the first bipartition will be equal to 60/(60+35), whereas the percentage of GTs supporting the second bipartition will be 35/(60+35). The reason that the number of GTs that support the third conflicting bipartition is not included is because we want IC to measure the magnitude of certainty conveyed by the two most prevalent bipartitions. This way, IC will equal zero when the two most prevalent bipartitions are equally prevalent (in this example that would be the case if bipartitions 1 and 2 were each supported by 42.5 GTs each).



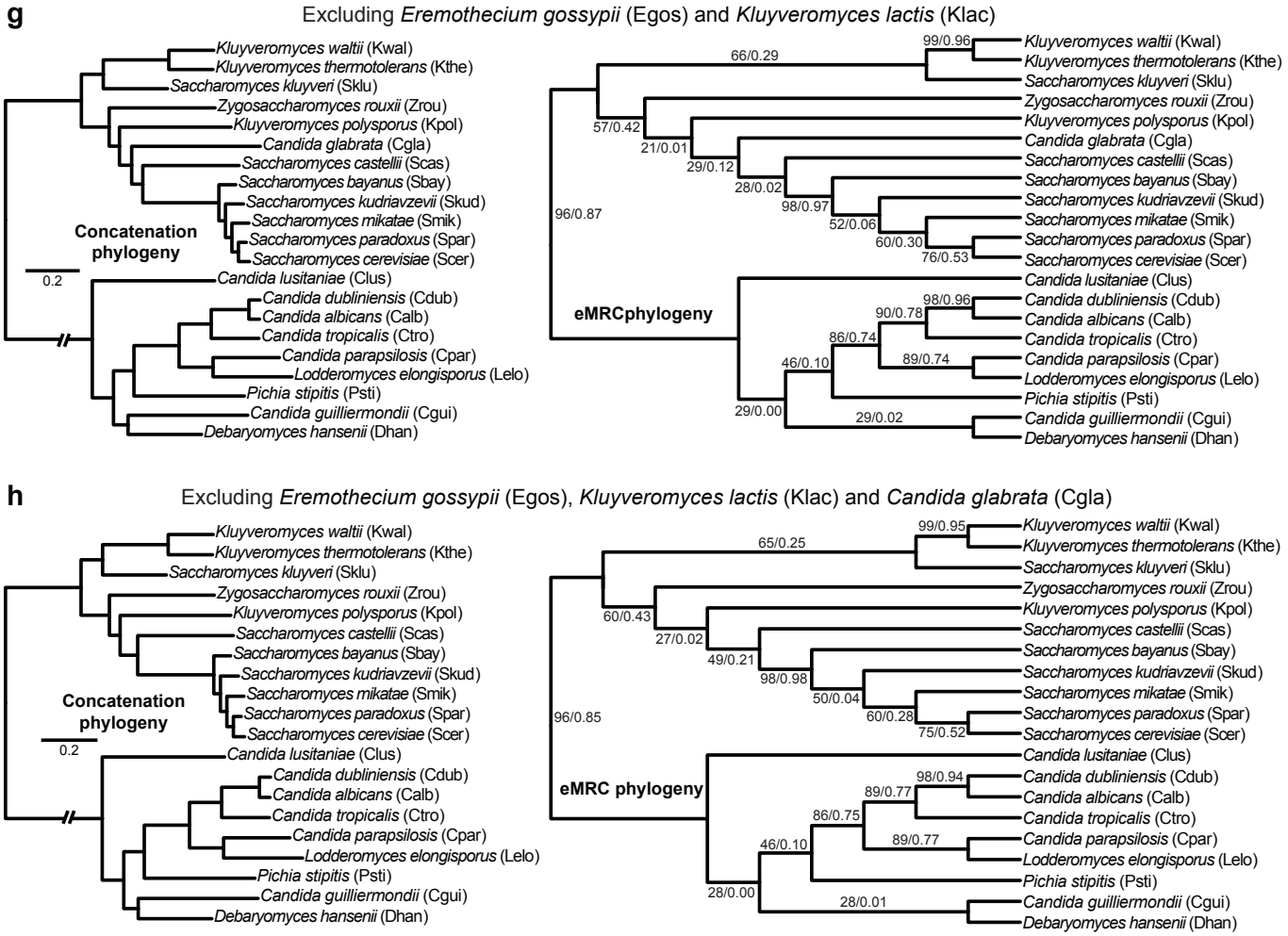
Supplementary Figure 3 | Removal of sites containing gaps or of poorly aligned genes does not significantly improve the yeast phylogeny inferred by concatenation and eMRC approaches. Each panel shows the yeast species phylogeny inferred from concatenation analysis (left panel) and from extended majority rule consensus (eMRC) analysis (right panel). All internodes of phylogenies inferred by concatenation received 100% bootstrap support unless otherwise indicated. Values near internodes of phylogenies inferred by eMRC analysis correspond to gene support frequency and internode certainty, respectively. **a**, Concatenation (left) and eMRC (right) phylogenies of all 1,070 genes following removal of all sites containing gaps. **b**, Concatenation (left) and eMRC (right) phylogenies of all 1,070 genes following removal of all sites where ≥50% of the character states are gaps. **c**, Concatenation (left) and eMRC (right) phylogenies of the 374 genes whose alignment length following removal of all gaps is ≥70% of the length of the original alignment.



Supplementary Figure 4 | Removal of one or more unstable or fast-evolving species has, if any, a minor and local effect on the yeast phylogeny inferred by concatenation and eMRC approaches. Each panel shows the yeast species phylogeny inferred from concatenation analysis (left panel) and from extended majority rule consensus (eMRC) analysis (right panel) following removal of one or more unstable or fast-evolving species from the analysis. All internodes of phylogenies inferred by concatenation received 100% bootstrap support unless otherwise indicated. Values near internodes of phylogenies inferred by eMRC analysis correspond to gene support frequency and internode certainty, respectively. **a**, Concatenation (left) and eMRC (right) phylogenies of all 1,070 genes following the removal of the unstable taxon *Candida lusitaniae*. **b**, Concatenation (left) and eMRC (right) phylogenies of all 1,070 genes following the removal of the fast-evolving and unstable taxon *Kluyveromyces polysporus*. **c**, Concatenation (left) and eMRC (right) phylogenies of all 1,070 genes following the removal of the fast-evolving and unstable taxon *Candida glabrata*.

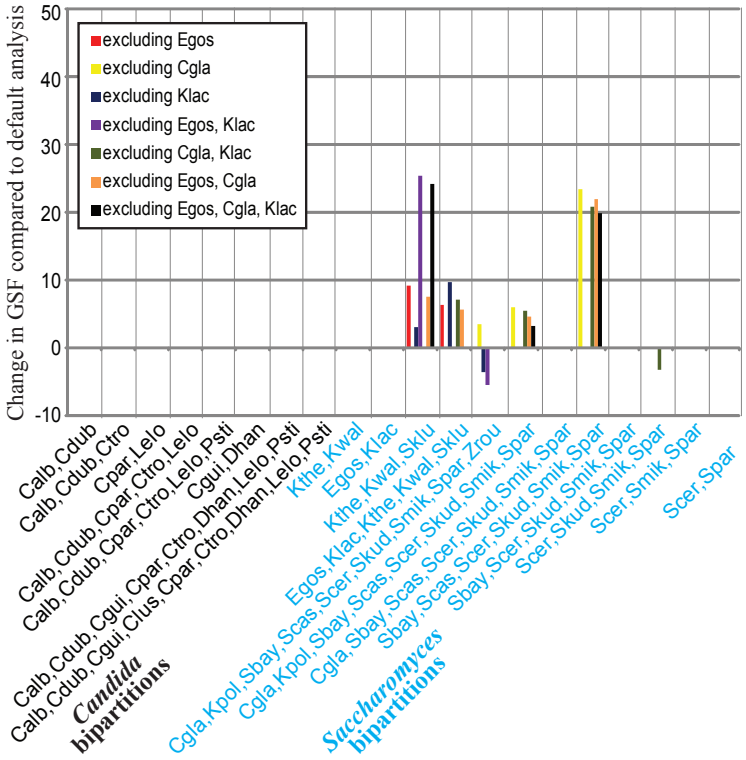


Supplementary Figure 4 | ...continued. d, Concatenation (left) and eMRC (right) phylogenies of all 1,070 genes following the removal of the unstable taxon *Saccharomyces castellii*. **e**, Concatenation (left) and eMRC (right) phylogenies of all 1,070 genes following the removal of the fast-evolving and unstable taxon *Eremothecium gossypii*. **f**, Concatenation (left) and eMRC (right) phylogenies of all 1,070 genes following the removal of the fast-evolving and unstable taxon *Kluyveromyces lactis*.

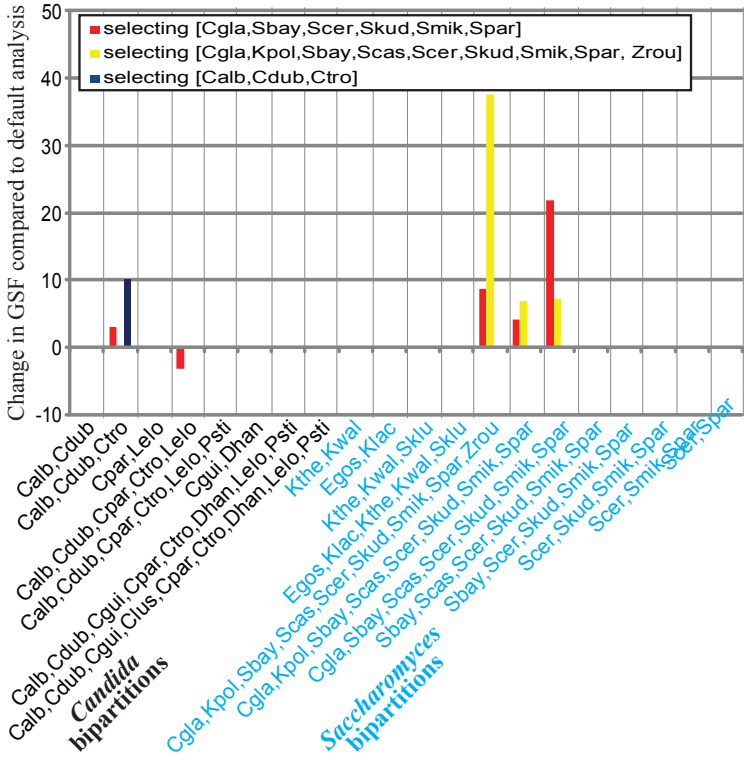


Supplementary Figure 4 | ...continued. g, Concatenation (left) and eMRC (right) phylogenies of all 1,070 genes following the removal of both *Eremothecium gossypii* and *Kluyveromyces lactis*. **h**, Concatenation (left) and eMRC (right) phylogenies of all 1,070 genes following the removal of *Eremothecium gossypii*, *Kluyveromyces lactis* and *Candida glabrata*.

a Change in GSF when removing fast-evolving and unstable species

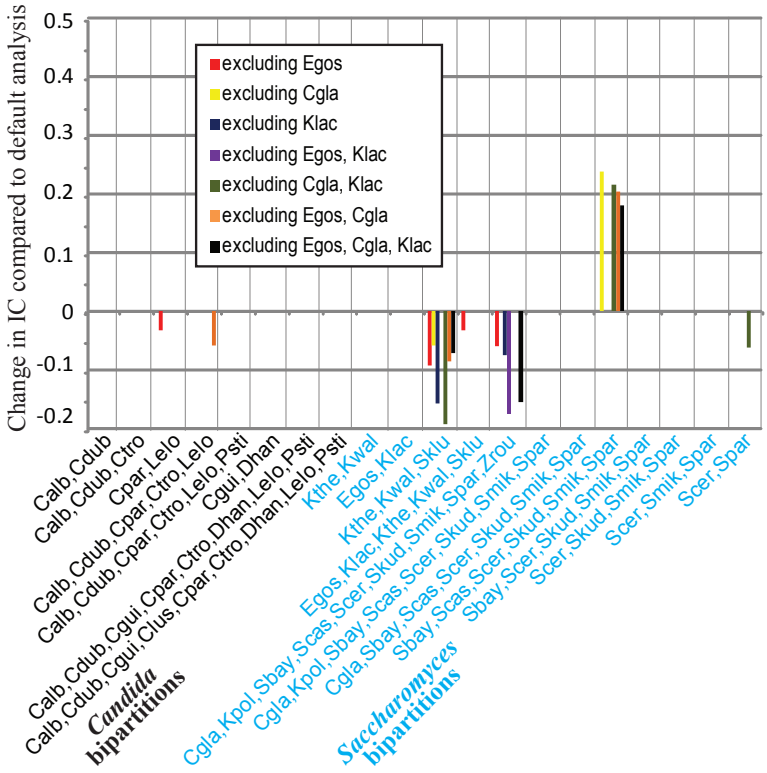


b Change in GSF when using only genes that support specific, likely correct, bipartitions

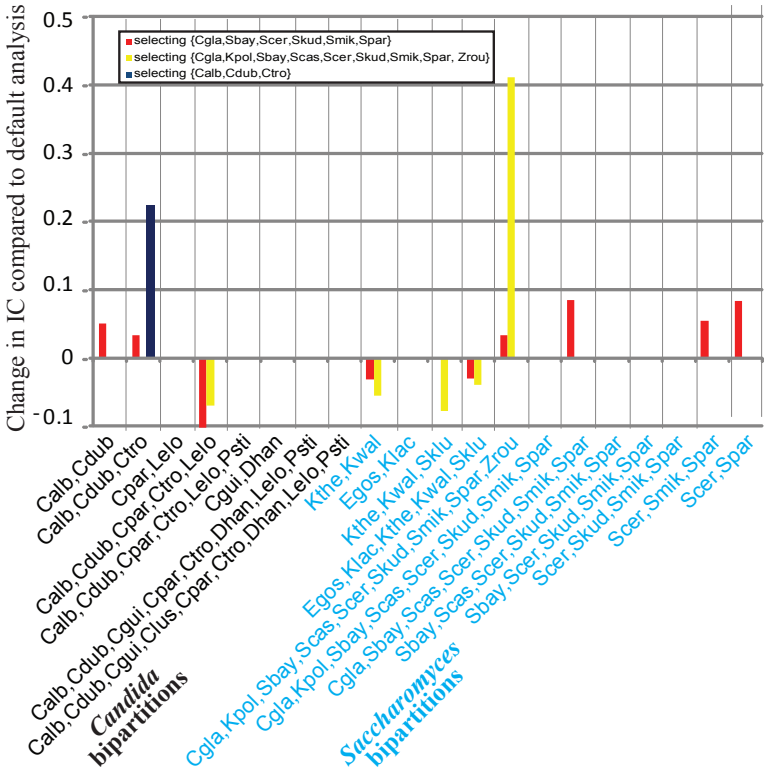


Supplementary Figure 5 | Removal of fast-evolving and unstable species or the exclusive use of genes that recover specific bipartitions has a minor and typically local effect on GSF values of internodes of the yeast phylogeny. The X-axis shows the 20 bipartitions present in the yeast phylogeny suggested by concatenation analysis and the Y-axis the percent change in gene support frequency (GSF) observed for each bipartition between the treatment and the default analysis. Only GSF changes ≥3% are shown. **a**, The individual or combined removal of *E. gossypii* (Egos), *K. lactis* (Klac), and *C. glabrata* (Cgla), three of the fastest evolving species as well as of those whose phylogenetic position is most unstable from the dataset has a minor and local effect on the GSF of neighboring internodes. Note that removal of species in the *Saccharomyces* lineage does not change the GSF of bipartitions in the *Candida* lineage. **b**, The selection of genes whose individual topologies recover well-established bipartitions of the yeast phylogeny has a minor effect on the GSF of internodes of the yeast phylogeny. Note that the [*C. albicans*, *C. dubliniensis*, *C. tropicalis*] (abbreviated [Calb, Cdub, Ctro]) bipartition has 90% GSF in the eMRC phylogeny reconstructed from the 1,070 individual gene trees, the [*C. glabrata*, *K. polysporus*, S. *bayanus*, *S. castellii*, *S. cerevisiae*, *S. kudriavzevii*, *S. mikatae*, *S. paradoxus*, *Z. rouxii*] (abbreviated [Zrou, Kpol, Cgla, Sbay, Skud, Smik, Scer, Spar]) bipartition has 62% GSF, and the [*C. glabrata*, *S. bayanus*, *S. cerevisiae*, *S. kudriavzevii*, *S. mikatae*, *S. paradoxus*] (abbreviated [Cgla, Sbay, Skud, Smik, Scer, Spar]) bipartition has 20% GSF. This last bipartition does not appear in the eMRC phylogeny but several independent rare genomic changes strongly suggest that it is the correct one.

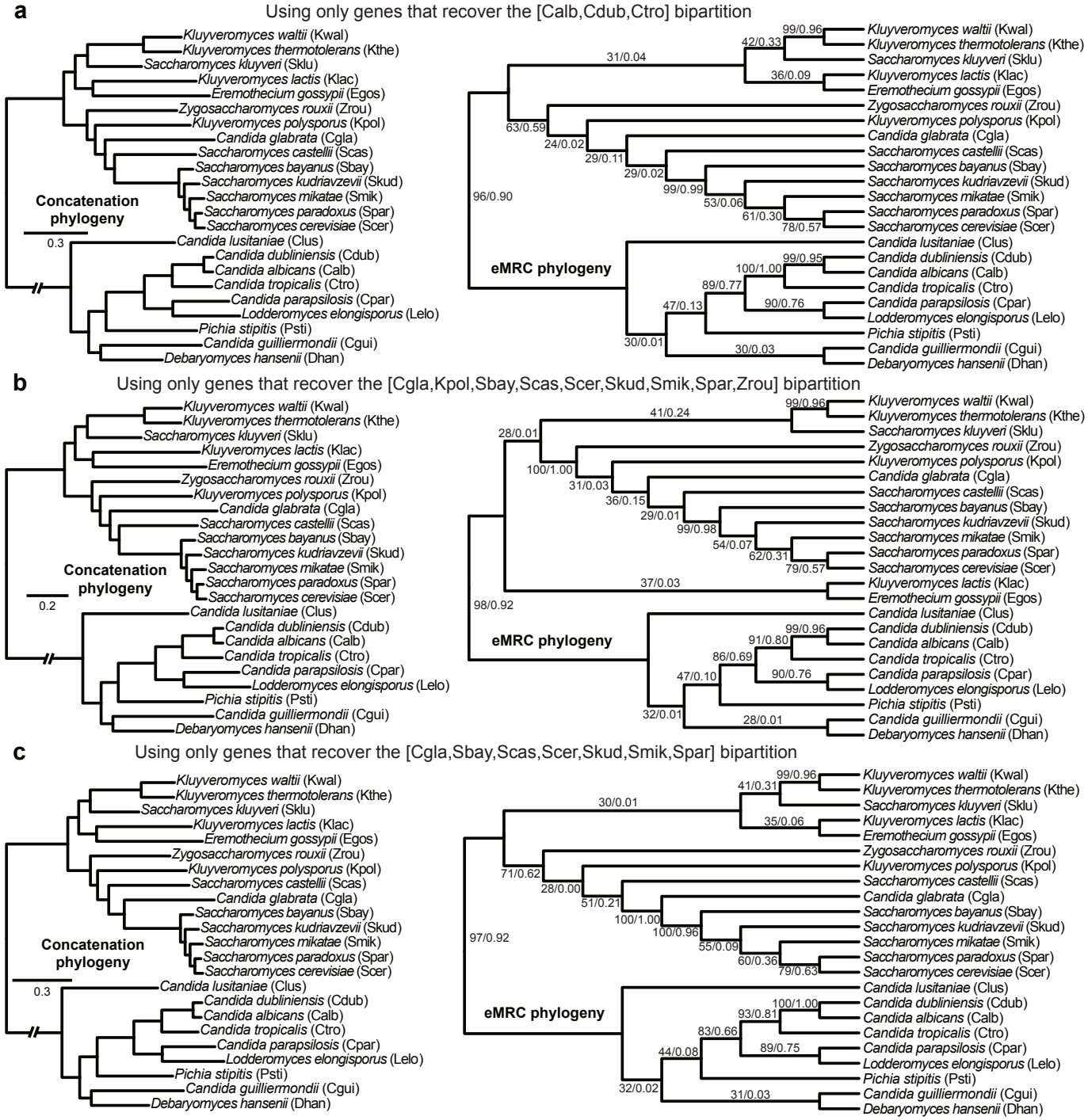
a Change in IC when removing fast-evolving and unstable species



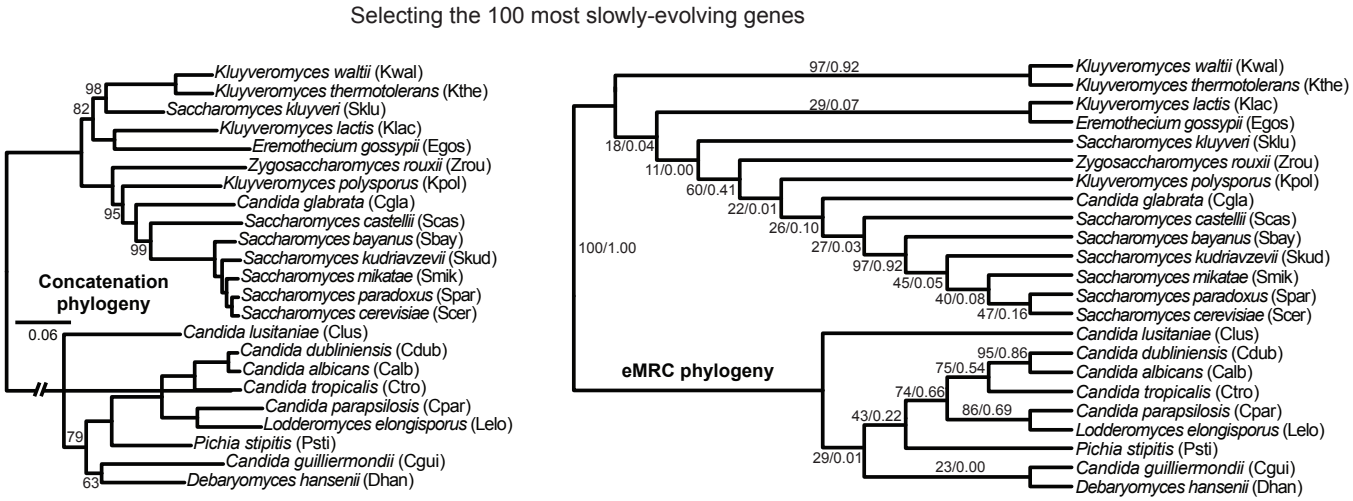
b Change in IC when using only genes that support specific, likely correct, bipartitions



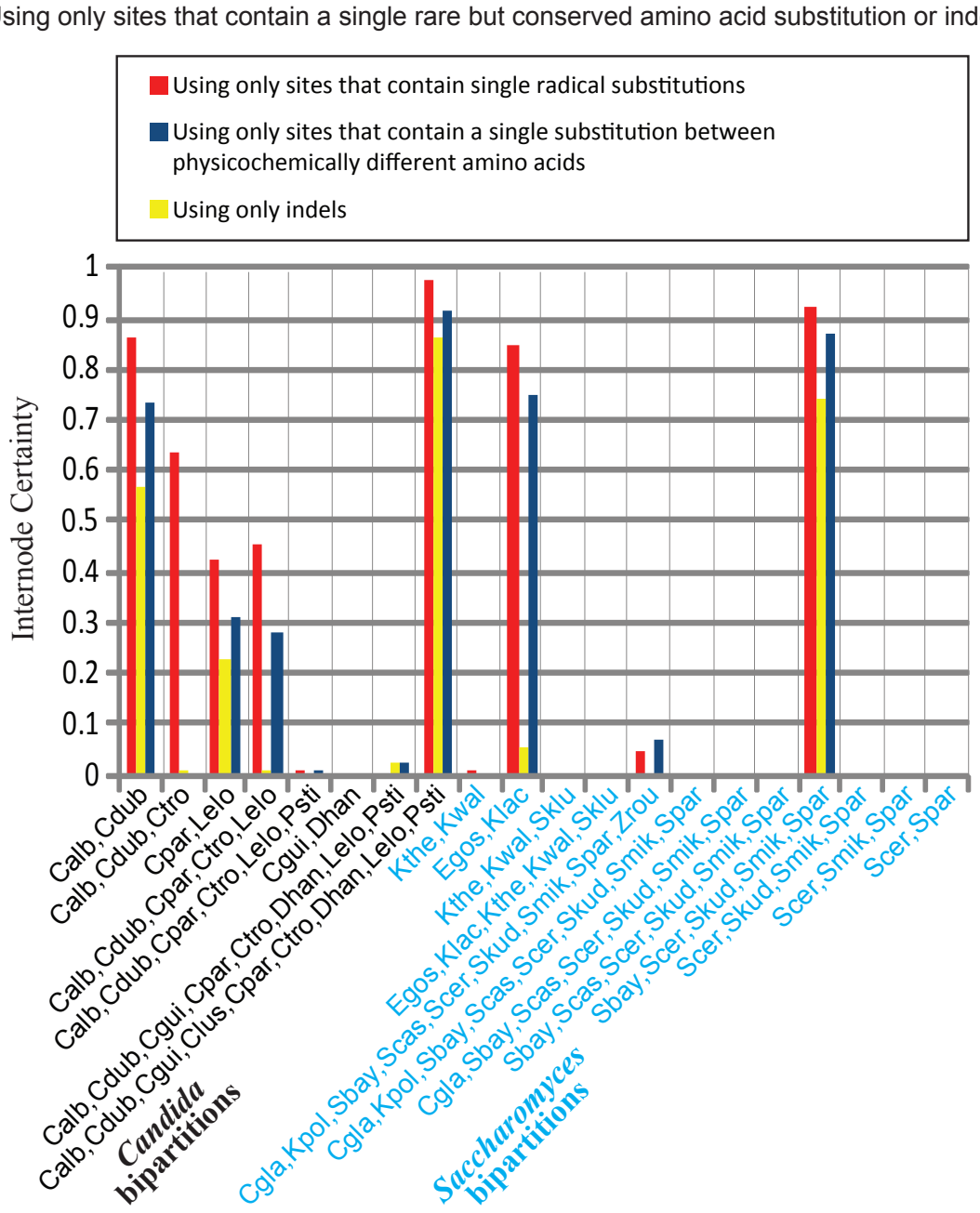
Supplementary Figure 6 | Removal of fast-evolving and unstable species or the exclusive use of genes that recover specific bipartitions has a minor and typically local effect on IC values of internodes of the yeast phylogeny. The X-axis shows the 20 bipartitions present in the yeast phylogeny suggested by concatenation analysis and the Y-axis the percent change in Internode Certainty (IC) observed for each bipartition between the treatment (removal of fast-evolving and unstable species or of genes that fail to recover specific clades) and the default analysis (all species and genes included). Only GSF changes ≥3% are shown. **a**, The individual or combined removal of *E. gossypii* (Egos), *K. lactis* (Klac), and *C. glabrata* (Cgla), three of the fastest evolving species as well as of those whose phylogenetic position is most unstable from the dataset has a minor and local effect on the IC of neighboring internodes. **b**, The selection of genes whose individual topologies recover well-established bipartitions of the yeast phylogeny has a minor effect on the IC of internodes of the yeast phylogeny. Note that the [*C. albicans*, *C. dubliniensis*, *C. tropicalis*] (abbreviated [Calb, Cdub, Ctro]) bipartition has 90% GSF in the extended majority rule consensus (eMRC) phylogeny reconstructed from the 1,070 individual gene trees, the [*C. glabrata*, *K. polysporus*, *S. bayanus*, *S. castellii*, *S. cerevisiae*, *S. kudriavzevii*, *S. mikatae*, *S. paradoxus*, *Z. rouxii*] (abbreviated [Zrou, Kpol, Cgla, Sbay, Skud, Smik, Scer, Spar]) bipartition has 62% GSF, and the [*C. glabrata*, *S. bayanus*, *S. cerevisiae*, *S. kudriavzevii*, *S. mikatae*, *S. paradoxus*] (abbreviated [Cgla, Sbay, Skud, Smik, Scer, Spar]) bipartition has 20% GSF. This last bipartition does not appear in the eMRC phylogeny but, as discussed in the main text, several independent rare genomic changes strongly suggest that it is the correct one.



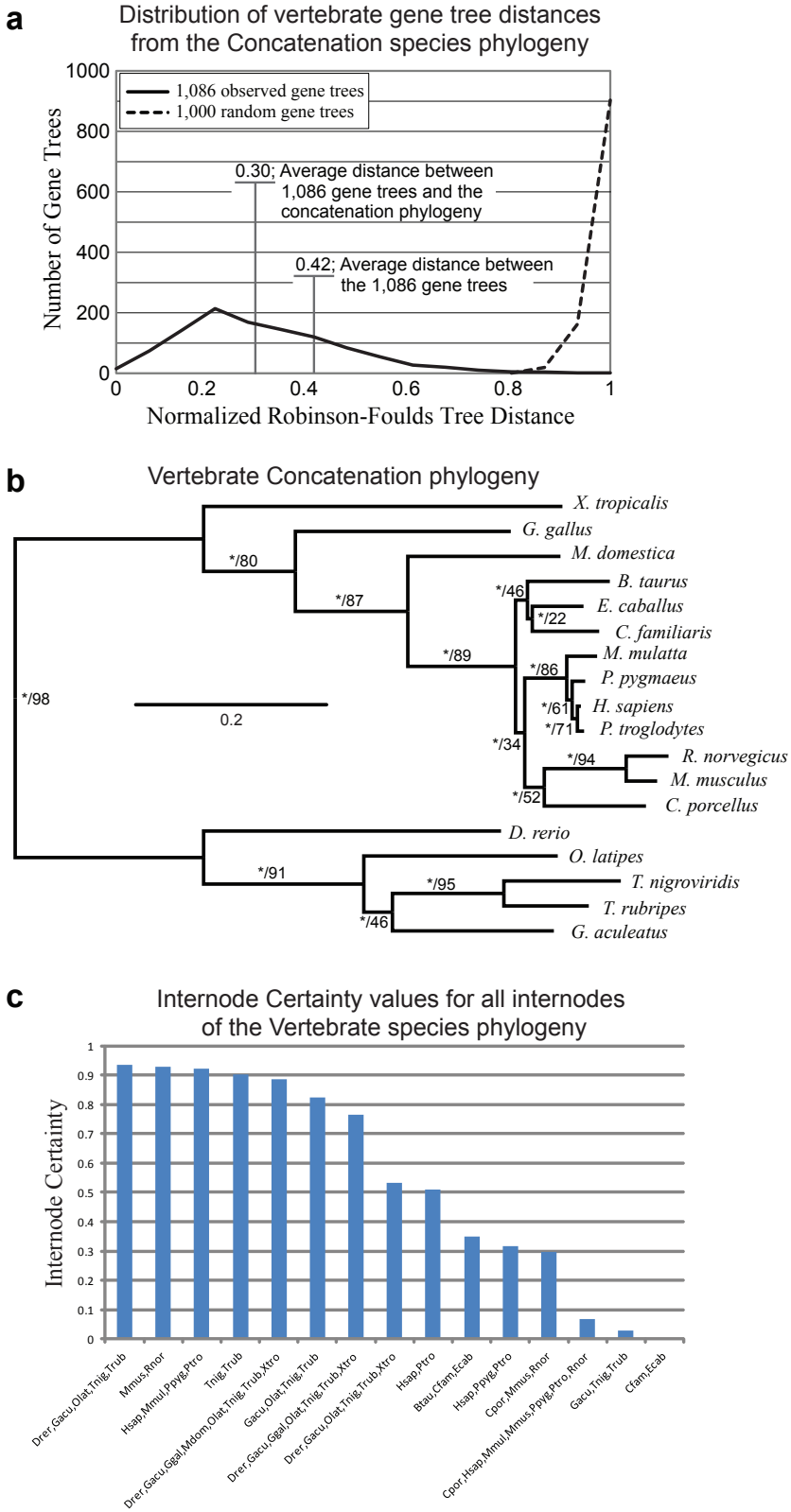
Supplementary Figure 7 | Selection of genes that support specific bipartitions has, if any, a minor and local effect on the yeast phylogeny inferred by concatenation and eMRC approaches. Each panel shows the yeast species phylogeny inferred from concatenation analysis (left panel) and from extended majority rule consensus (eMRC) analysis (right panel) following the selection and use of genes that recover specific bipartitions. All internodes of phylogenies inferred by concatenation received 100% bootstrap support unless otherwise indicated. Values near internodes of phylogenies inferred by eMRC analysis correspond to gene support frequency and internode certainty, respectively. **a**, Concatenation (left) and eMRC (right) phylogenies using only the genes that recover the [*C. albicans*, *C. dubliniensis*, *C. tropicalis*] (abbreviated [Calb, Cdub, Ctro]) bipartition. **b**, Concatenation (left) and eMRC (right) phylogenies using only the genes that recover the [*C. glabrata*, *K. polysporus*, *S. bayanus*, *S. castellii*, *S. cerevisiae*, *S. kudriavzevii*, *S. mikatae*, *S. paradoxus*, *Z. rouxii*] (abbreviated [Zrou, Kpol, Cgla, Sbay, Skud, Smik, Scer, Spar]) bipartition. **c**, Concatenation (left) and eMRC (right) phylogenies using only the genes that recover the [*C. glabrata*, *S. bayanus*, *S. cerevisiae*, *S. kudriavzevii*, *S. mikatae*, *S. paradoxus*] (abbreviated [Cgla, Sbay, Skud, Smik, Scer, Spar]) bipartition. This last bipartition does not appear in the eMRC phylogeny but, as discussed in the main text, several independent rare genomic changes strongly suggest that it is the correct one.



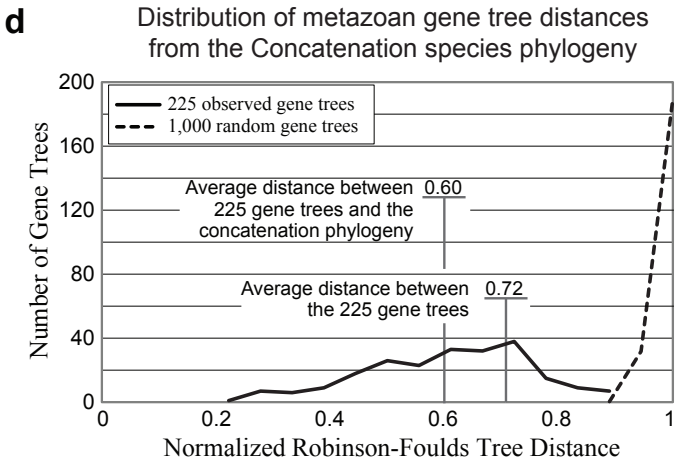
Supplementary Figure 8 | Selection of the 100 slowest-evolving genes has a large, negative effect on GSF and IC values of internodes of the yeast phylogeny inferred by concatenation and eMRC approaches. Each panel shows the yeast species phylogeny inferred from concatenation analysis (left panel) and from extended majority rule consensus (eMRC) analysis (right panel) following the selection and use of the 100 slowest-evolving genes. All internodes of phylogenies inferred by concatenation received 100% bootstrap support unless otherwise indicated. Values near internodes of phylogenies inferred by eMRC analysis correspond to gene support frequency and internode certainty, respectively.

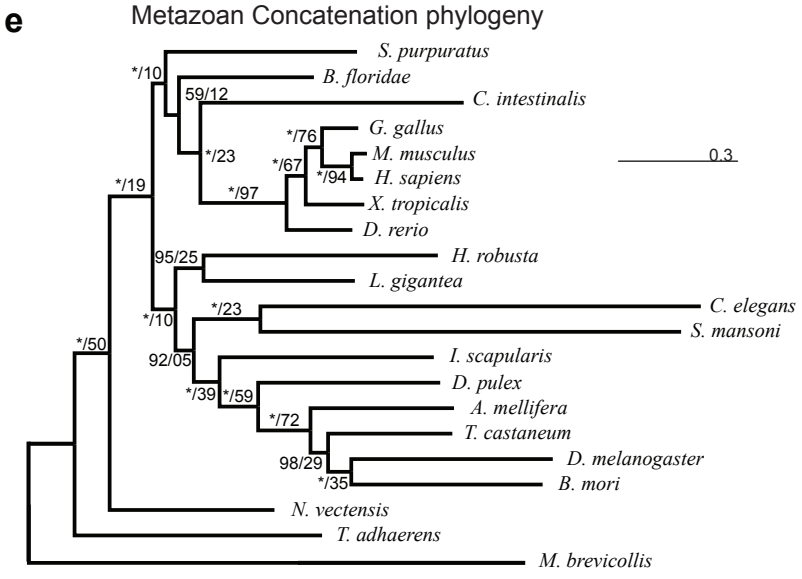


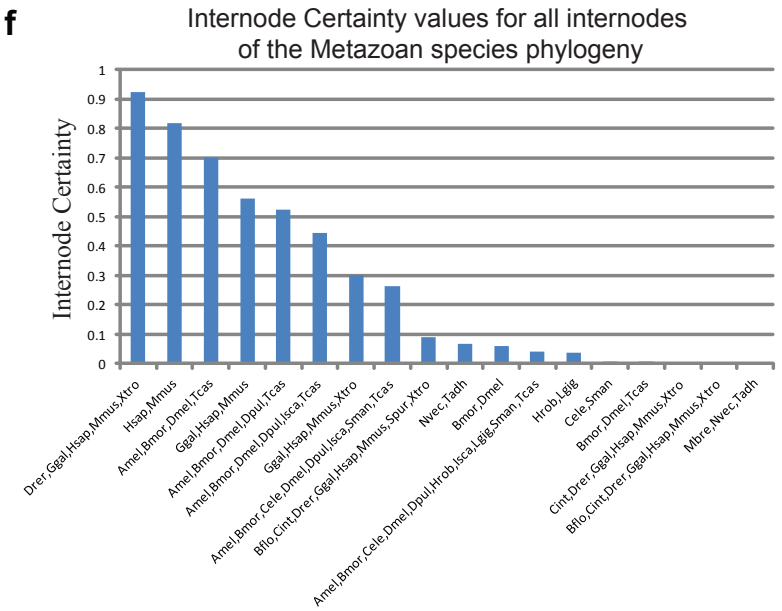
Supplementary Figure 9 | Selection of sites that contain a single rare but conserved amino acid substitution or indels has a large, negative effect on GSF and IC values of internodes of the yeast phylogeny. The X-axis shows the 20 bipartitions present in the yeast phylogeny suggested by concatenation analysis and the Y-axis the percent change in Internode Certainty (IC) observed for each bipartition between the treatment (selection of specific sites or indels) and the default analysis (all sites included). Only GSF changes ≥3% are shown. The red bars correspond to changes in IC when using only the 20,289 sites that contain single radical substitutions (defined as a substitution with a blosum62 matrix score ≤–3), the blue bars correspond to changes in IC when using only the 4,075 sites that contain a single substitution between amino acids that differ radically in their physicochemical properties, and the yellow bars correspond to changes in IC when using only the 2,474 characters which mark the presence / absence of a single indel that spans 7 or more aa among the 23 yeast species.



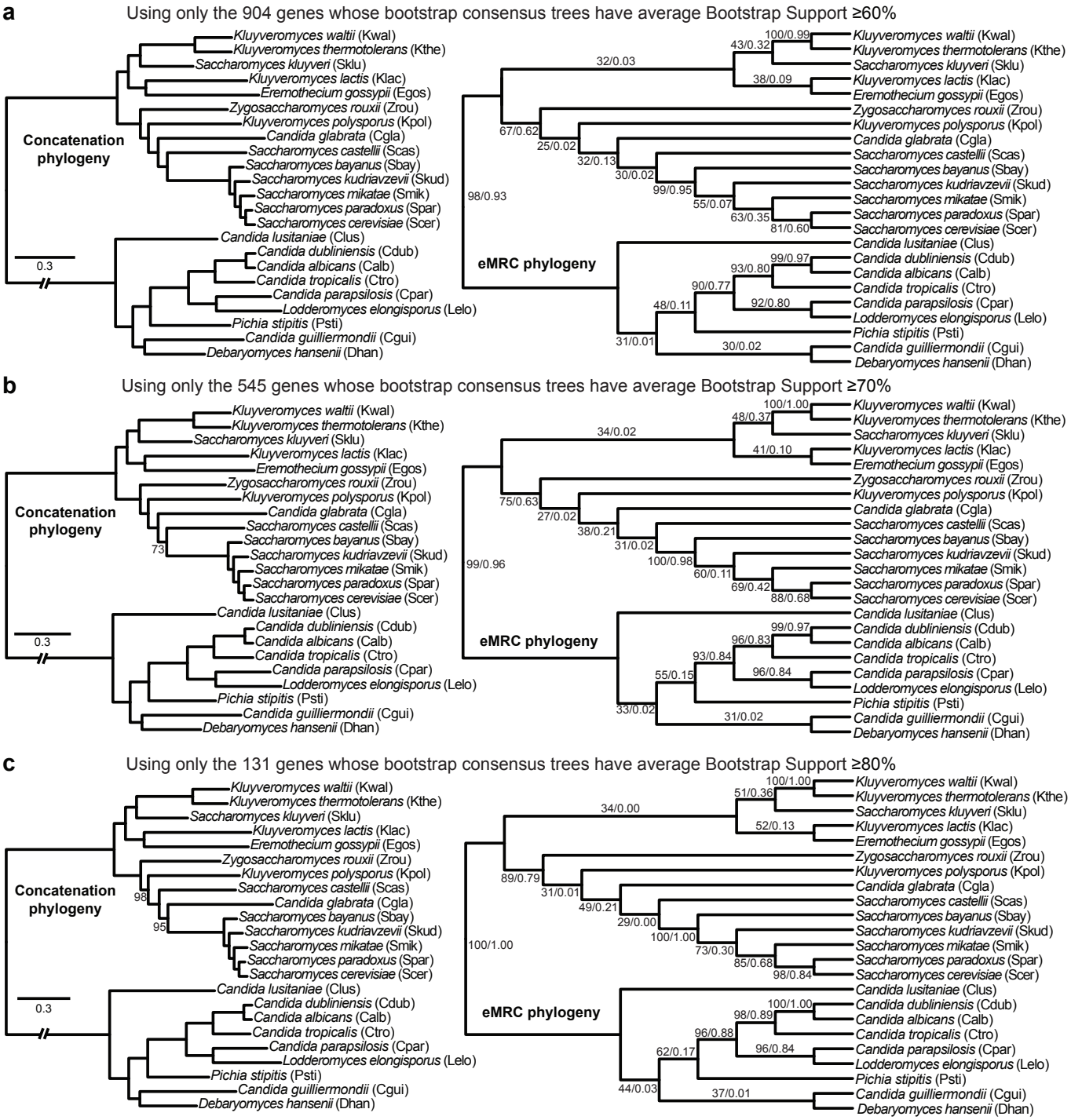
Supplementary Figure 10 | High levels of incongruence in Vertebrate and Metazoan phylogenomic datasets despite the inference of highly supported phylogenies by concatenation analysis. **a**, The distribution of the agreement between the bipartitions present in the 1,086 individual gene trees (GTs) and the vertebrate concatenation phylogeny, as well as the distribution of the agreement between the bipartitions present in 1,000 randomly generated trees of equal taxon number and the concatenation phylogeny, measured using the normalized Robinson-Foulds tree distance. The average tree distances between the 1,086 GTs and the concatenation phylogeny as well as between the 1,086 GTs with each other are also shown. **b**, The vertebrate species phylogeny recovered from concatenation analysis of 1,086 genes using maximum likelihood. The extended majority rule consensus (eMRC) phylogeny is topologically identical to the concatenation phylogeny. Values near internodes correspond to bootstrap support and gene support frequency (GSF), respectively. Asterisks (*) denote internodes that received 100% bootstrap support by the concatenation analysis. **c**, The distribution of Internode Certainty (IC) values for all internodes of the vertebrate species phylogeny.



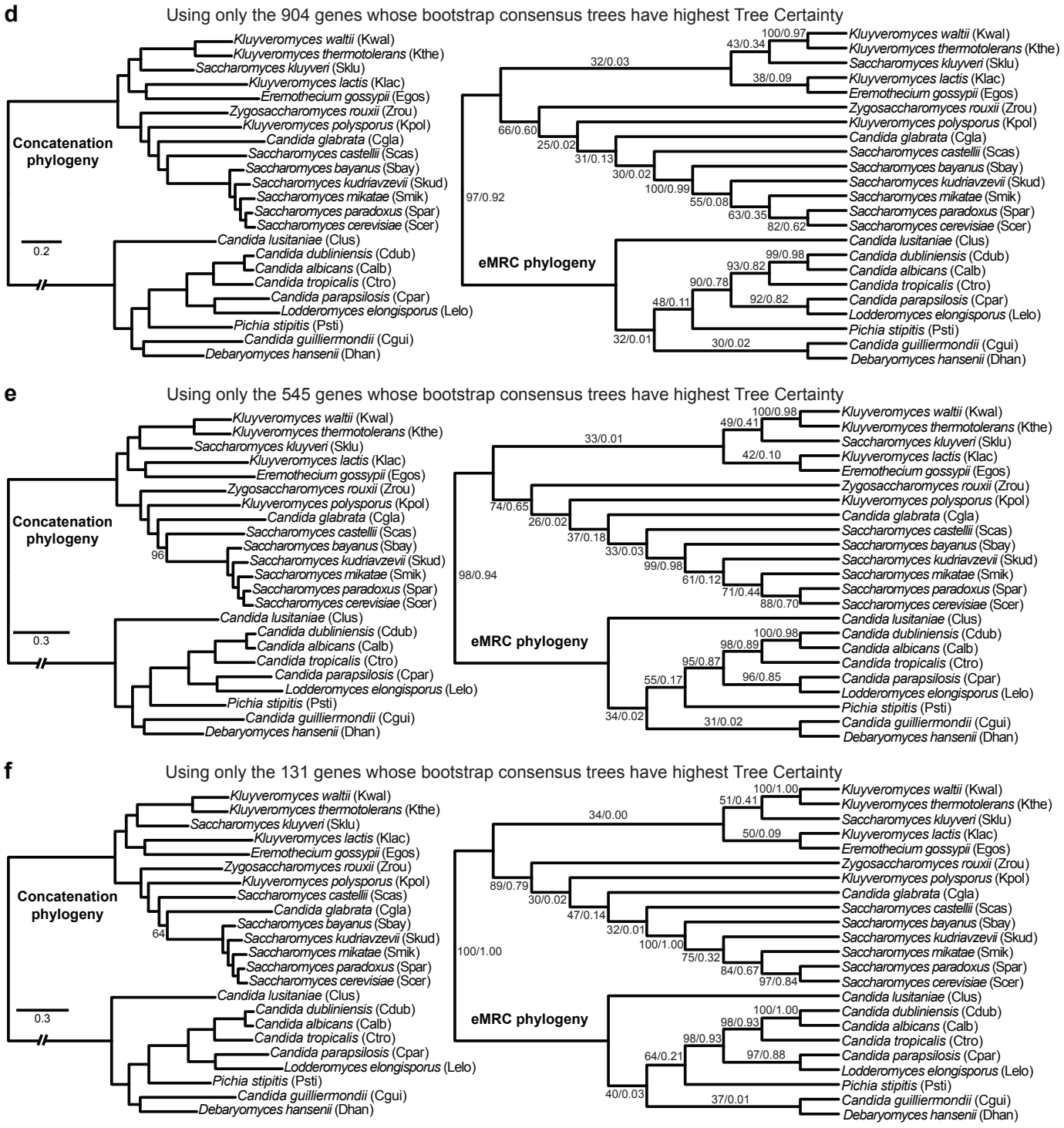




Supplementary Figure 10 | ...continued. d, The distribution of the agreement between the bipartitions present in the 225 individual GTs and the metazoan concatenation phylogeny, as well as the distribution of the agreement between the bipartitions present in 1,000 randomly generated trees of equal taxon number and the concatenation phylogeny, measured using the normalized Robinson-Foulds tree distance. The average tree distances between the 225 GTs and the concatenation phylogeny as well as between the 225 GTs with each other are also shown. e, The metazoan species phylogeny recovered from concatenation analysis of 225 genes using maximum likelihood. The eMRC phylogeny is topologically identical to the concatenation phylogeny. Values near internodes correspond to bootstrap support and gene support frequency, respectively. Asterisks (*) denote internodes that received 100% bootstrap support by the concatenation analysis. f, The distribution of IC values for all internodes of the metazoan species phylogeny. Note that GSF and IC values indicate the existence of numerous internodes in the vertebrate and especially in the metazoan phylogeny that are supported by a small percentage of gene trees and have very small or zero IC values.

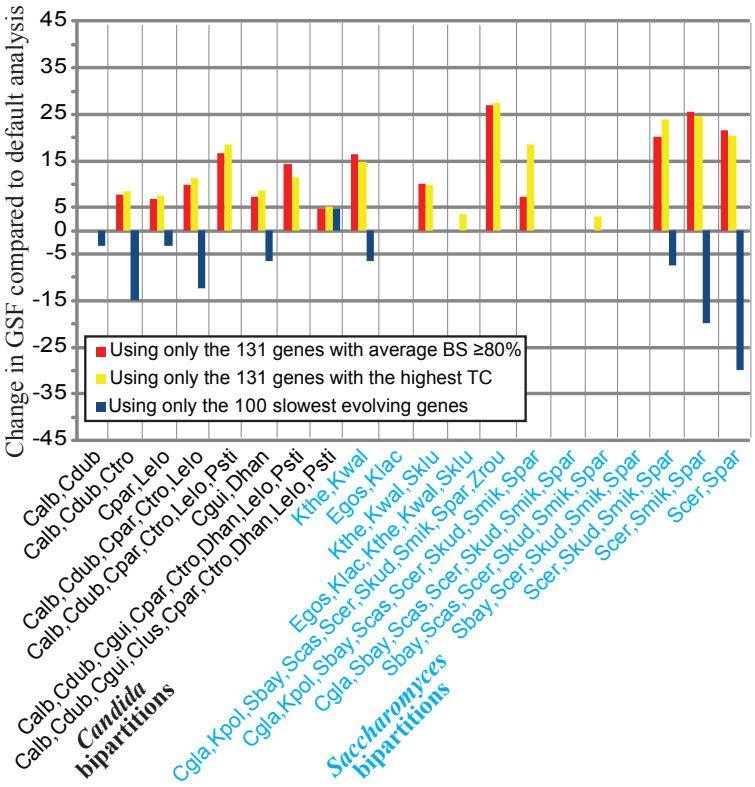


Supplementary Figure 11 | Selection of genes whose bootstrap consensus trees have high average Bootstrap Support (avBS) or Tree Certainty (TC) has a large, positive effect on GSF and IC values of internodes of the yeast phylogeny inferred by concatenation and eMRC approaches. Each panel shows the yeast species phylogeny inferred from concatenation analysis (left panel) and from extended majority rule consensus (eMRC) analysis (right panel) following the selection of genes whose trees have high average bootstrap support (BS) or Tree Certainty (TC). All internodes of phylogenies inferred by concatenation received 100% bootstrap support unless otherwise indicated. Values near internodes of phylogenies inferred by eMRC analysis correspond to gene support frequency and internode certainty, respectively. **a**, Concatenation (left) and eMRC (right) phylogenies of the 904 genes whose gene trees have average BS ≥60%. **b**, Concatenation (left) and eMRC (right) phylogenies of the 545 genes whose gene trees have average BS ≥70%. **c**, Concatenation (left) and eMRC (right) phylogenies of the 131 genes whose gene trees have average BS ≥80%.

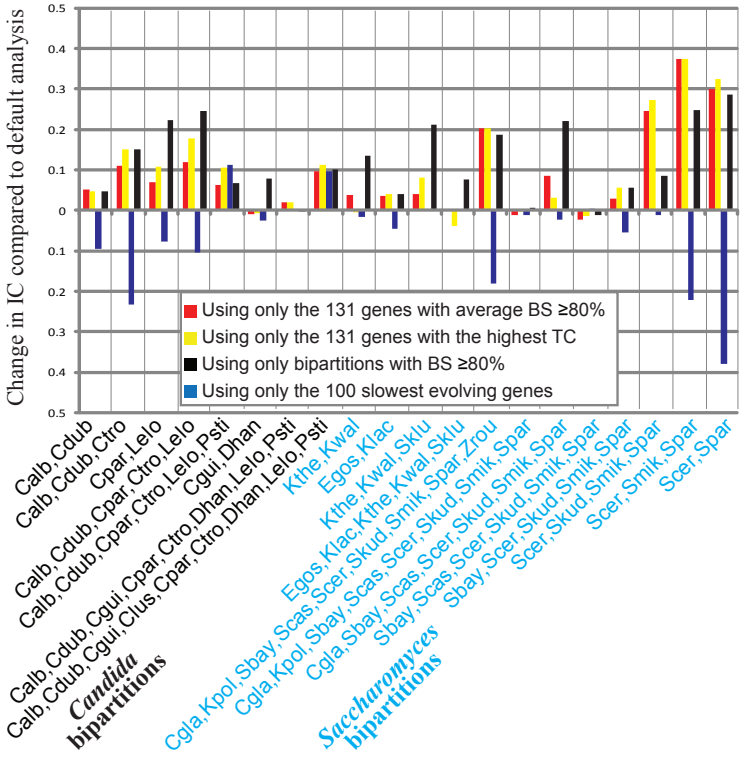


Supplementary Figure 11 | ...continued. d, Concatenation (left) and eMRC (right) phylogenies of the 904 genes with the highest TC. **e**, Concatenation (left) and eMRC (right) phylogenies of the 545 genes with the highest TC. **f**, Concatenation (left) and eMRC (right) phylogenies of the 131 genes with the highest TC.

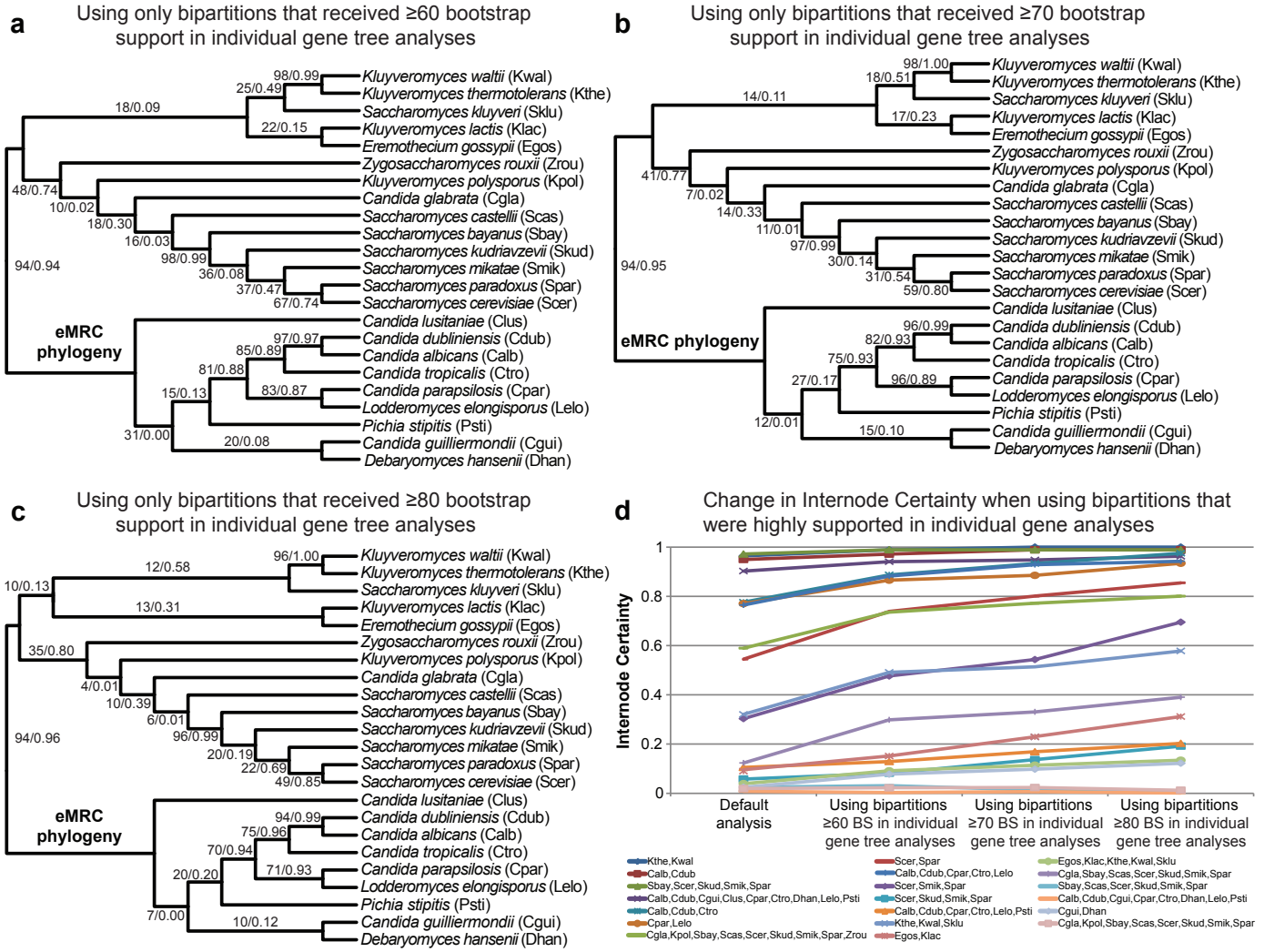
a Change in Gene Support Frequency for highly supported genes or slow evolving genes



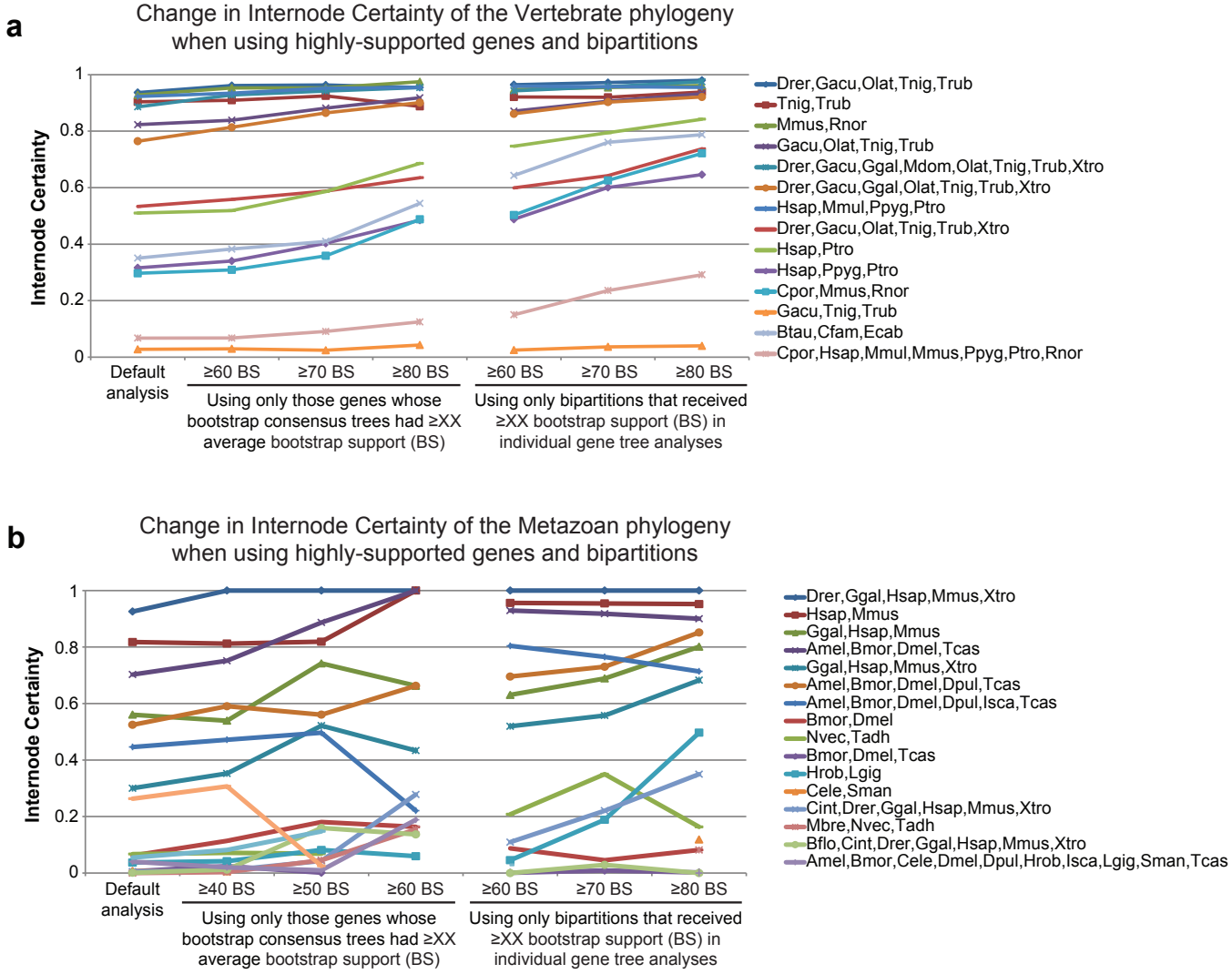
b Change in Internode Certainty for highly supported genes, bipartitions or slow evolving genes



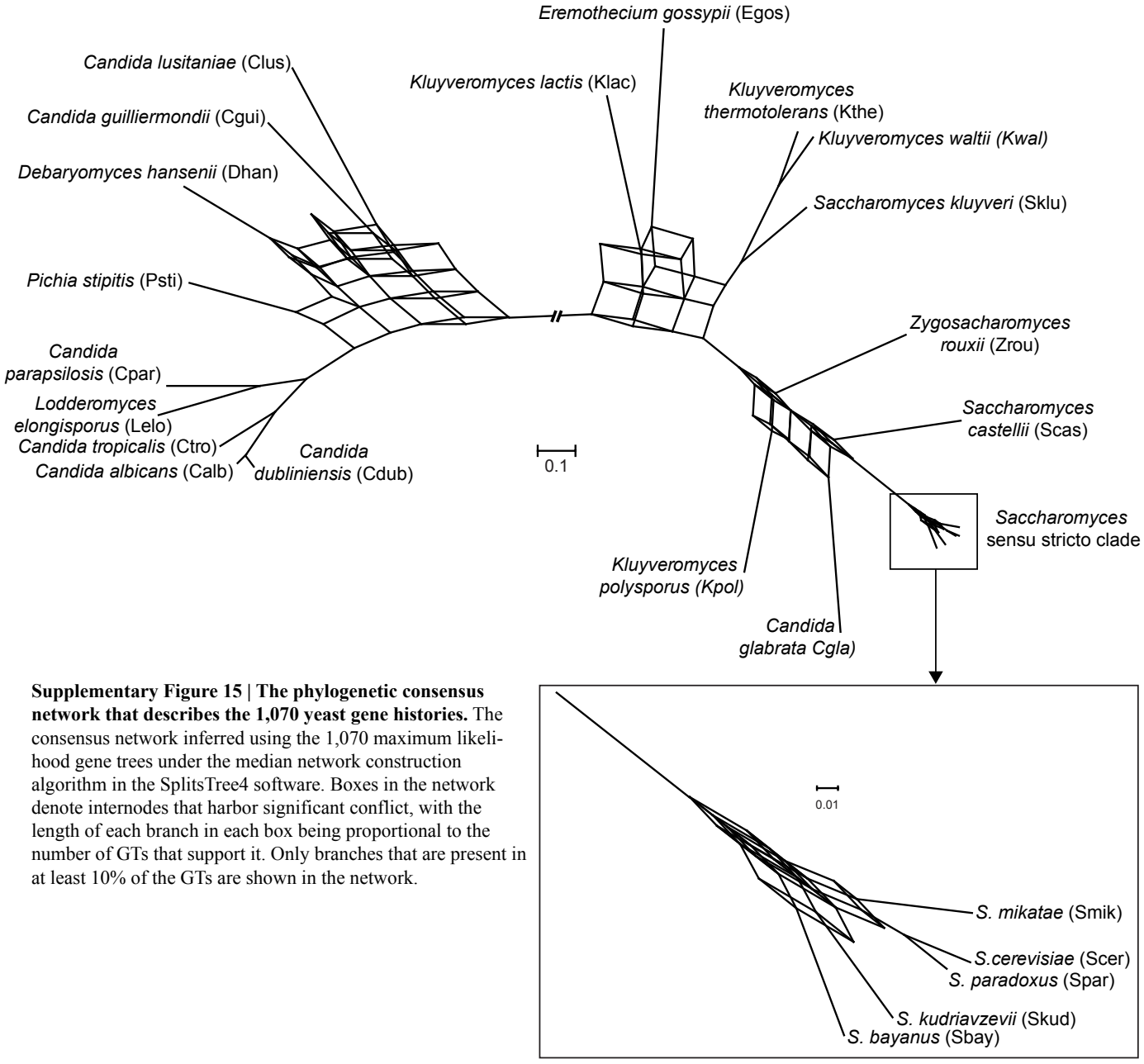
Supplementary Figure 12 | Selecting highly supported genes or bipartitions has a large, positive effect on GSF and IC values of internodes of the yeast phylogeny. The X-axis shows the 20 bipartitions present in the yeast phylogeny suggested by concatenation analysis and the Y-axis the percent change in Gene Support Frequency (GSF) and Internode Certainty (IC) observed for each bipartition between the treatment (selection of highly supported genes or internodes) and the default analysis. Only GSF changes ≥3% and IC changes ≥0.03 are shown. The red bars correspond to changes in IC when using only the 131 genes with average bootstrap support ≥80%, the yellow bars correspond to changes in IC when using only the 131 genes with the highest Tree Certainty, the black bars correspond to changes in IC when using only those bipartitions found in the bootstrap consensus trees of individual genes that had bootstrap support ≥80%, and the blue bars correspond to changes in IC when using only the 100 slowest-evolving genes. **a**, Change in GSF for highly supported genes or slow evolving genes. **b**, Change in IC for highly supported genes, bipartitions or slow evolving genes.



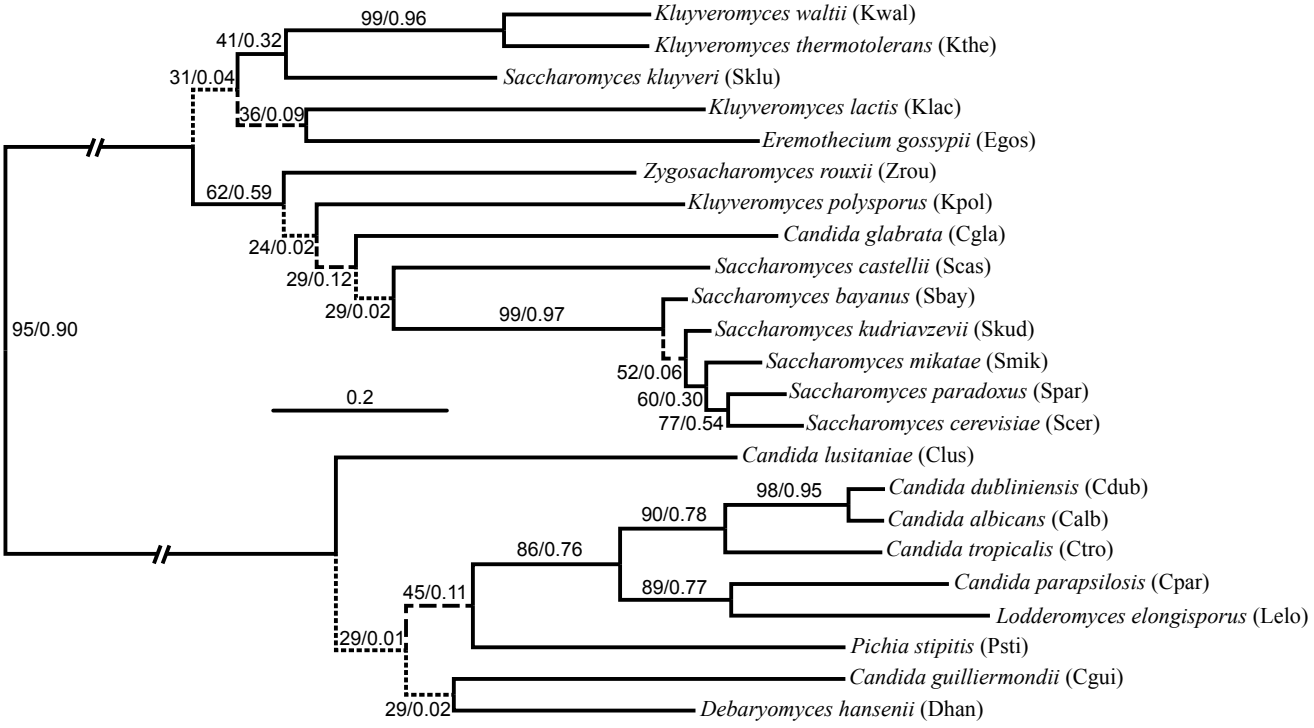
Supplementary Figure 13 | Selection of highly supported bipartitions from the bootstrap consensus trees of individual genes has a large, positive effect on the IC values of internodes of the yeast phylogeny inferred by the eMRC approach. The first three panels show the yeast species phylogeny inferred from extended majority rule consensus (eMRC) analysis following the selection of bipartitions that had high bootstrap support (BS) in the bootstrap consensus trees of individual genes. Values near internodes correspond to the percentage of bootstrap consensus trees of individual genes in which this specific bipartition received high BS and to internode certainty (IC), respectively. **a**, The eMRC phylogeny inferred from selecting bipartitions that had BS ≥60% in individual gene analyses. **b**, The eMRC phylogeny inferred from selecting bipartitions that had BS ≥70% in individual gene analyses. **c**, The eMRC phylogeny inferred from selecting bipartitions that had BS ≥80% in individual gene analyses. **d**, Plot that illustrates the change in IC of internodes relative to the values obtained in the default analysis associated with the use of bipartitions that had high bootstrap support (BS) in the bootstrap consensus trees of individual genes. Each line of different color depicts the IC value obtained for a given internode in the default analysis (Fig. 1a), when using only bipartitions that had BS ≥60%, BS ≥70%, and BS ≥80%.



Supplementary Figure 14 | Selection of highly supported genes and bipartitions has a large, positive effect on IC values of internodes of the vertebrate and metazoan phylogenies. a, Plot that illustrates the change in IC of internodes of the vertebrate phylogeny relative to the values obtained in the default analysis associated with the use of genes whose bootstrap consensus trees have high average bootstrap support (BS) or with the use of bipartitions that had high BS in the bootstrap consensus trees of individual genes. Each line of different color depicts the IC value obtained for a given internode in the default analysis (Supplementary Fig. S10b), when using only genes with average BS ≥60%, BS ≥70%, and BS ≥80%, as well as when using only bipartitions that had BS ≥60%, BS ≥70%, and BS ≥ 80%. **b**, Plot that illustrates the change in IC of internodes of the metazoan phylogeny relative to the values obtained in the default analysis associated with the use of genes whose bootstrap consensus trees have high average bootstrap support (BS) or with the use of bipartitions that had high BS in the bootstrap consensus trees of individual genes. Each line of different color depicts the IC value obtained for a given internode in the default analysis (Supplementary Fig. S10e), when using only genes with average BS ≥40%, BS ≥50%, and BS ≥60%, as well as when using only bipartitions that had BS ≥60%, BS ≥70%, and BS ≥80%. The slight differences in the pattern of change of IC in the metazoan dataset relative to the other two datasets is likely to be due to the much smaller number of metazoan genes in the dataset and their lower average bootstrap support, which will result in smaller sets of genes and bipartitions with strong phylogenetic signal.

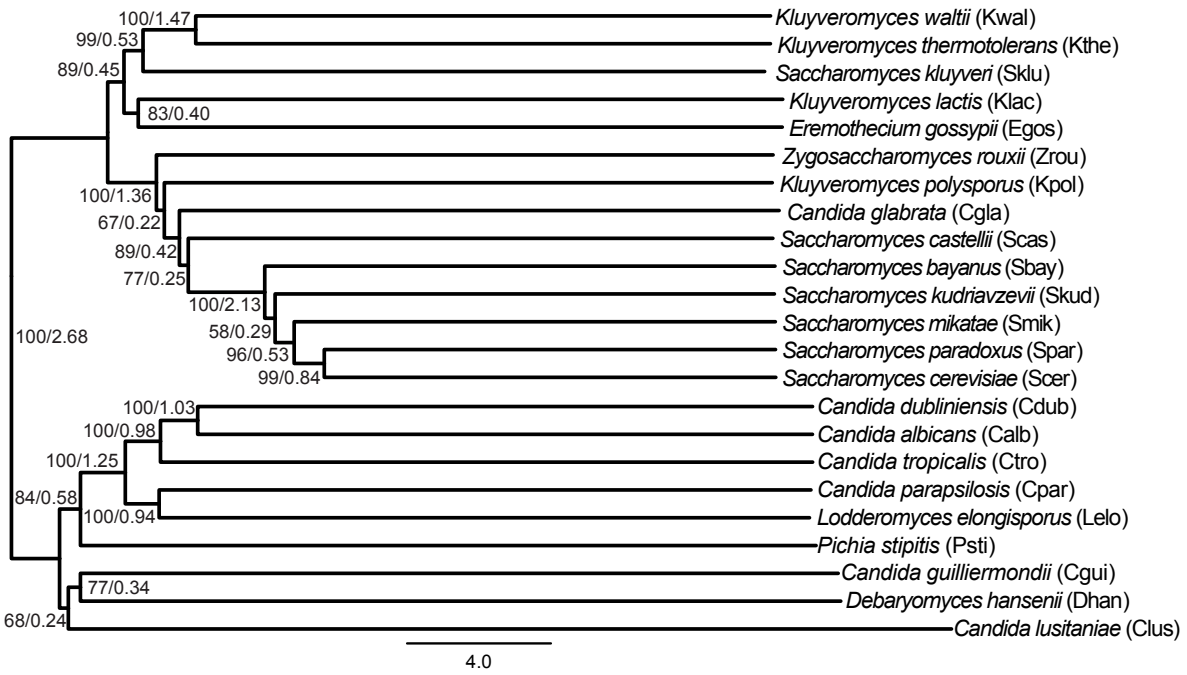


Supplementary Figure 15 | The phylogenetic consensus network that describes the 1,070 yeast gene histories. The consensus network inferred using the 1,070 maximum likelihood gene trees under the median network construction algorithm in the SplitsTree4 software. Boxes in the network denote internodes that harbor significant conflict, with the length of each branch in each box being proportional to the number of GTs that support it. Only branches that are present in at least 10% of the GTs are shown in the network.

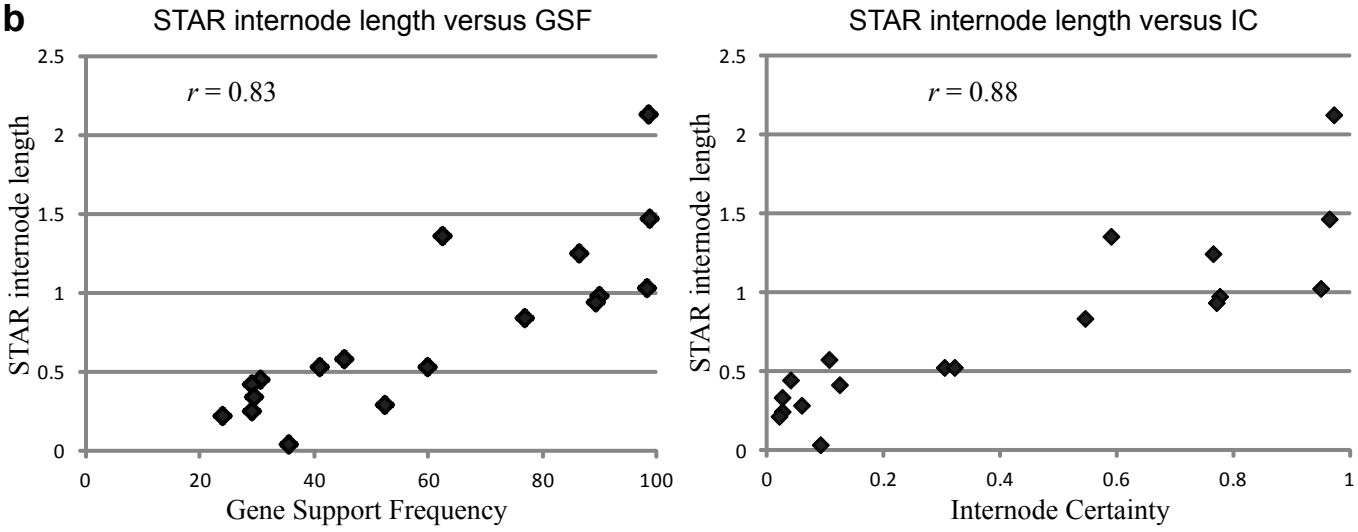


Supplementary Figure 16 | Supported, weakly supported, and unresolved internodes in the yeast phylogeny. Values near internodes correspond to gene support frequency and internode certainty, respectively calculated from the 1,070 yeast gene histories. Note that the validity of certain internodes marked as "unresolved" is supported by independent data (e.g., rare genomic changes).

a The yeast species phylogeny inferred using the STAR species tree method



b



Supplementary Figure 17 | The yeast phylogeny inferred using a "species tree" method that accounts for variation between the 1,070 gene histories is highly supported and has extremely short internodes whose coalescent unit lengths are highly correlated with gene support frequency and internode certainty values. Using the 1,070 gene dataset, we inferred a yeast species phylogeny under the coalescent model and average ranks of gene coalescence times, as implemented in the STAR species tree method. **a**, The yeast species phylogeny under the coalescent. Values near internodes correspond to bootstrap support and internode length in coalescence units, respectively. The inferred topology is identical to the phylogeny shown in Figure 1a, except with respect to the placement of *Candida lusitaniae*. **b**, The lengths of internodes in the phylogeny inferred using the STAR species tree method, measured in average coalescent units, is highly correlated with internodes' Gene Support Frequency (left panel) and Internode Certainty (right panel) values. The strength of each correlation is indicated by *r*, Pearson's correlation coefficient.