Supplemental figure legends for **Haase *et al.***

Figure S1. Genome-scale maximum likelihood phylogeny. See methods section for details on phylogenetic analyses.

Figure S2. Genome-scale internode certainty cladogram. See methods section for details on phylogenetic analyses.

Figure S3. Distribution of the structure of *GAL* gene clusters.

Both cluster structure and growth characteristics are mapped onto the relative divergence timetree. Growth on galactose is indicated by the colored squares next to each species (green=blue, yellow=variable, red=negative). Asterisks next to certain species names indicated either a new genome sequence published here (\*\*) or an additional genome sequence from a recent study (\*) (Shen *et al.* 2018), including *Nadsonia fulvescens* var. *fulvescens*. To ensure phenotyping could be performed on sequenced strains, we also sequenced the genomes of the taxonomic type strains for eight species and report those *GAL* clusters here (^). The syntenic structure of the *GAL* genes are displayed to the right of the growth characteristics for each species. The structure of the *Nadsonia fulvescens* var. *elongata* cluster is shown in Figure S4.

Figure S4. Syntenic *GAL* clusters between diverse lineages. Numbers correspond to positions in each scaffold or contig. Shaded regions indicate regions of synteny between clusters.

Figure S5. Alignment of the *GAL10* genes of *N. fulvescens* var. *fulvescens* and *N. fulvescens* var. *elongota*. Genes were aligned using MAFFT v 7.409 using -- auto. Likely inactivating mutations are shown in various colors: mutation of the start codon in orange, frameshift mutations in blue, in-frame nonsense mutations in red, and insertions in green. One in-frame deletion is shown in purple.

Figure S6. Percent identities of *GAL* genes as calculated by the comparisons shown in Figure 3B.

(**A**) Raw percent identy values are shown for all groups and color-coded by each protein. Comparisons are as described for Figure 3B.

(**B** – **D**) Similarity scores for each group for the label protein. Comparisons are as described for Figure 3B.

Figure S7. Gene tree of *GAL1* genes. See methods section for details on phylogenetic analyses.

Figure S8. Gene tree of *GAL7* genes. See methods section for details on phylogenetic analyses.

Figure S9. Gene tree of *GAL10 genes*. See methods section for details on phylogenetic analyses.

Figure S10. Concatenated gene tree of the *GAL*actose enzymatic gene cluster. See methods section for details on phylogenetic analyses.

Figure S11. Phylogenetic constraint trees used in the Approximately Unbiased tests.

(A-D) Phylogenetic constraint trees of the concatenated *GAL1*, *GAL7*, and *GAL10* genes used in the Approximately Unbiased test in Figure 4 (the same topology was used for each individual gene analysis). (A) Constrained to follow the species tree, except for the 3 HGT candidate lineages, (B) additionally constrained for only *Brettanomyces* to fit the species topology, (C) additionally constrained for only *Nadsonia* to fit the species topology, and (D) additionally constrained for only *Wickerhamomyces* to fit the species topology.