Evolution of a novel chimeric maltotriose transporter in *Saccharomyces eubayanus* from parent proteins unable to perform this function.

Short title: Evolution of maltotriose consumption via a novel chimeric protein.

EmilyClare P. Baker[1,2] and Chris Todd Hittinger[1,2,3]*

[1]Laboratory of Genetics, Genome Center of Wisconsin, Wisconsin Energy Institute, J. F. Crow Institute for the Study of Evolution, University of Wisconsin-Madison, Madison, WI 53706, USA

[2]Microbiology Doctoral Training Program, University of Wisconsin-Madison, Madison WI, 53706

[3]DOE Great Lakes Bioenergy Research Center, University of Wisconsin-Madison, Madison, WI 53706, USA

*corresponding author

EB helped conceive of this project, designed and performed all experiments and analyses, and wrote this manuscript with input and edits from CTH.

CTH conceived of this project, helped write the manuscript, supervised all work, and provided financial support.

Abstract

At the molecular level, the evolution of new traits can be broadly divided between changes in gene expression and changes in protein structure. For proteins, the evolution of novel functions is generally thought to proceed through sequential point mutations or recombination of whole functional units. In *Saccharomyces*, the uptake of the sugar maltotriose into the cell is the primary limiting factor in its utilization, but maltotriose transporters are relatively rare, except in brewing strains. No known wild strains of *Saccharomyces eubayanus*, the cold-tolerant parent of hybrid lager-brewing yeasts (*Saccharomyces cerevisiae x S. eubayanus*), are able to consume maltotriose, which limits their ability to fully ferment malt extract. In one strain of *S. eubayanus*, we found a gene closely related to a known maltotriose transporter and were able to confer maltotriose consumption by overexpressing this gene or by passaging the strain on maltose. Even so, most wild strains of *S. eubayanus* lack native maltotriose transporters. To determine how this rare trait could evolve in naive genetic backgrounds, we performed an adaptive evolution experiment for maltotriose consumption, which yielded a single strain of *S. eubayanus* able to grow on maltotriose. We mapped the causative locus to a gene encoding a novel chimeric transporter that was formed by an ectopic recombination event between two genes encoding transporters that are unable to import maltotriose. In contrast to classic models of the evolution of novel protein functions, the recombination breakpoints occurred within functional domains. Thus, the ability of the new protein to carry maltotriose was likely acquired through epistatic interactions between independently evolved substitutions. By acquiring multiple mutations at once, the transporter rapidly gained a novel function, while bypassing potentially deleterious intermediate steps. This study provides an illuminating example of how recombination between paralogs can establish novel interactions among substitutions to create adaptive functions.

Author summary

Hybrids of the yeasts *Saccharomyces cerevisiae* and *Saccharomyces eubayanus* (lager-brewing yeasts) dominate the modern brewing industry. *S. cerevisiae*, also known as baker's yeast, is well-known for its role in industry and scientific research. Less well recognized is *S. eubayanus*, which was only discovered as a pure species in 2011. While most lager-brewing yeasts rapidly and completely utilize the important brewing sugar maltotriose, no strain of *S. eubayanus* isolated to date is known to do so. Despite being unable to consume maltotriose, we identified one strain of *S. eubayanus* carrying a gene for a functional maltotriose transporter, although most strains lack this gene. During an adaptive evolution experiment, a strain of *S. eubayanus* without native maltotriose transporters evolved the ability to grow on maltotriose. Maltotriose consumption in the evolved strain resulted from a chimeric transporter that arose through recombination between genes encoding parent proteins that were unable to transport maltotriose. Traditionally, functional chimeric proteins are thought to evolve by recombining discrete functional domains or modules, but the breakpoints in the chimera studied here occurred within modular units of the protein. These results support the less well-recognized role of recombination between paralogous sequences in generating novel proteins with adaptive functions.

Introduction

Proteins with novel functions can arise through a variety of mechanisms (1). One of the best studied mechanisms is gene duplication, followed by divergence through sequential point mutations (1,2). While this method of new protein evolution is thought to be common, evolution through stepwise point mutations can be a slow and constrained process (3). In the mutational landscape separating the original protein from the derived protein, deleterious epistatic interactions, where multiple intermediate mutational steps interact to create fitness valleys, can make new functions difficult to access by successive point mutations. Mutational events that result in multiple amino acid changes at

once can help bridge fitness valleys and speed the evolution of new functionality (3,4). As a consequence of bypassing intermediate mutational steps, recombination can lead to intragenic reciprocal sign epistasis, where the new recombinant protein has a function not found in either parent protein (3,5). Ectopic gene conversion, which results in chimeric protein sequences, is one such rare class of mutational events that can rapidly lead to new protein sequences with novel functions (3).

Chimeric protein-coding sequences have been found to be an important mechanism by which proteins can evolve new functions (1,6–8). They have been implicated in the rapid radiation of multicellular animals (6) and in playing a role in both infectious and non-infectious diseases in humans (9–12). The *Drosophila* gene *jingwei* was one of the first chimeric genes to have both its recent origin and evolution characterized in depth (1). *jingwei* exemplifies many of the characteristics usually associated with chimeric proteins (1,6–8,13). Like most other chimeric proteins that have been described in eukaryotes, *jingwei* is a large multidomain protein that was constructed via the movement of whole functional units (domains), facilitated by intronic sequences, a process referred to as domain or exon shuffling. In most cases, even in the absence of intronic sequences, the recombination of these modules has been considered key to the evolution of functional chimeric proteins (1,14).

The exchange of complete, independently functional units is not the only method by which functional chimeric proteins can be generated. Recombination within functional domains also has the potential to create proteins with novel characteristics. Recombination breakpoints within domains can lead to functional proteins, even between non-homologous protein sequences (4,15,16). However, since functionally important structures are likely to be conserved between related proteins, the probability of recombination resulting in a functional protein is higher between homologous sequences where essential within-protein interactions are less likely to be disrupted (4,15,17). Theoretical work has suggested the potential of this sort of recombination to allow proteins to rapidly bypass fitness minima in the adaptive landscape separating two protein functions (3,4). Recombination between paralogous

sequences has also been shown to be selected for in natural populations, suggesting that such sequences can indeed produce functional proteins (15,18). In addition, recombination between paralogous sequences (DNA shuffling) has been used as an efficient way to engineer proteins with functions that are rare or difficult to evolve in natural settings (reviewed in (19–21)). For example, hexose sugar transporters in *Saccharomyces cerevisiae* were evolved for increased specificity to a pentose sugar, D-xylose (22), through DNA shuffling and selection for the ability to support growth on xylose.

Maltotriose, a trimer of glucose molecules, is the second most abundant fermentable sugar present in brewing wort (malt extract), but it is also the most difficult to ferment (23–26). Among budding yeasts of the genus *Saccharomyces*, such as *S. cerevisiae,* proteins that can transport maltotriose into the cell are relatively rare (27–31). Improving consumption of maltotriose by *Saccharomyces* yeasts is of general interest to the brewing community since a key consideration for any new brewing strain is its ability to rapidly and completely ferment all the sugars present in wort. Work on improving the direct uptake of maltotriose in brewing yeasts has focused on the expression of the limited set of known maltotriose transporters, either through adaptive evolution for increased expression (32,33), introducing maltotriose transporters into new strains through selective breeding (34–42), or by heterologous expression (33,40,43). These methods all rely on the presence of functional maltotriose transporters, either natively or heterologously expressed, and are limited by the number of strains and proteins that are known to be capable of transporting maltotriose. With the focus on known transporters, how new maltotriose transporters evolve is less well studied (44,45).

Recently, special interest has been given to the development of *Saccharomyces eubayanus,* a distant cold-tolerant relative of *S. cerevisiae,* for commercial brewing (35,36,46,47). As a hybrid with *S. cerevisiae, S. eubayanus* forms the industrially important lager-brewing yeasts (48), which account for more than 90% of the total beer market. So far, no strain of *S. eubayanus* isolated from nature has been

94 reported to consume maltotriose (36,49–53), despite evidence for the possible presence of functional
95 transporters in the *S. eubayanus* subgenome of industrial *S. cerevisiae* x *S. eubayanus* hybrids (i.e. lager-
96 brewing yeasts) (28,54–58).

97 In the present study, we characterize the native *MALT* genes found in *S. eubayanus* for their
98 ability to enable the transport of maltotriose and confirm the presence of maltotriose transporters in
99 one strain of *S. eubayanus*, despite its inability to consume maltotriose. We also describe a novel
100 chimeric maltotriose transporter that resulted from the adaptive evolution of *S. eubayanus* for
101 maltotriose consumption. This new maltotriose transporter was formed through a partial ectopic gene
102 conversion event between two *MALT* genes. Interestingly, the parent proteins that produced the
103 chimera were unable to transport maltotriose themselves. In addition, the breakpoints of the chimeric
104 region do not demarcate clearly defined functional domains, suggesting that epistatic interactions
105 between novel residue combinations, rather than domain swapping, is responsible for the new function.
106 Overall, this study reports the first known maltotriose transporters in *S. eubayanus* and the first strains
107 of this species that are able to consume maltotriose. In addition, by characterizing one of the few
108 chimeric proteins that have been described where recombination occurred naturally within functional
109 modules (without being specifically targeted for engineering by DNA shuffling or mutagenesis), we
110 provide insight into how proteins can evolve novel adaptive functions through rare genetic events.

111 ## Results/Discussion

112 ### Maltotriose transporters in *S. eubayanus*

113 In the type strain of *S.* eubayanus, four genes, designated *MALT1-4,* have been identified as
114 having homology to genes encoding known maltose transporters (*MALT* genes) (54,59). Because *MALT2*
115 and *MALT4* are predicted to encode identical amino acid sequences (see Materials and Methods), we
116 refer to these genes jointly as *MALT2/4*. To determine if they could enable maltotriose transport, Malt1,
117 Malt2/4, and Malt3 were individually overexpressed using an inducible promoter in yHRVM108, a strain

of *S. eubayanus* isolated from North Carolina that is unable to grow on maltotriose and, unlike other strains of *S. eubayanus*, grows sluggishly on maltose. None of these genes were able to confer growth on maltotriose when overexpressed (Table 1).

Table 1. Heterologous expression of *S. eubayanus MALT* genes.

Strain	Background	Transporter	Initial OD	Day 3	Day 6
yHRVM108*	North Carolinian strain	-	0.16 (+/-0.05)	0.39 (+/-0.02)	0.48 (+/-0.01)
yHRVM108	North Carolinian strain	-	0.12 (+/-0.03)	0.47 (+/-0.00)	0.46 (+/-0.03)
yHEB1870	yHRVM108	*MALT1*	0.13 (+/-0.03)	0.43 (+/-0.04)	0.58 (+/-0.04)
yHEB1877	yHRVM108	*MALT2/4*	0.11 (+/-0.00)	0.39 (+/-0.01)	0.57 (+/-0.02)
yHEB1872	yHRVM108	*MALT3*	0.13 (+/-0.01)	0.41 (+/-0.00)	0.62 (+/-0.5)
yHEB1883	yHRVM108	*ncAGT1*	0.11 (+/-0.01)	0.54 (+/-0.07)	1.34 (+/-0.10)
yHEB1884	yHRVM108	*lgAGT1*	0.10 (+/-0.00)	0.42 (+/-0.07)	0.94 (+/-0.09)

Growth on SC + 2% maltotriose (98% pure) of strains expressing *MALT* genes on a doxycycline-inducible plasmid. N = 3, standard deviation in parentheses. * Control grown in SC + 0.04% glucose + doxycycline to reflect the approximate amount of growth expected from contamination with other carbon sources when using 98% pure maltotriose.

Although none of the transporters found in the type strain of *S. eubayanus* were able to support growth on maltotriose, there is compelling evidence from lager-brewing yeasts for the existence of maltotriose transporters within the greater *S. eubayanus* population (28,54–57). Of particular interest are alleles of *AGT1*. Two versions of *AGT1* are present in the genomes of lager-brewing yeasts. One, which we call *scAGT1* (*S. cerevisiae-AGT1*), was donated by the *S. cerevisiae* parent of lager yeasts, and the other, which we call *lgAGT1* (lager-*AGT1*), has been proposed to be of *S. eubayanus* origin (55). Both *lgAGT1* and *scAGT1*, like other *AGT1* alleles, can transport both maltose and maltotriose (27,28,57,60–63). Thus far, full-length sequences closely related to this *lgAGT1* have not been described in any strain of *S. eubayanus* (36).

Strain CDFM21L.1 and a closely related strain isolated from North Carolina, yHRVM108, belong to the Holarctic subpopulation of *S. eubayanus* and are close relatives of the strains of *S. eubayanus* that

hybridized with *S. cerevisiae* to form lager-brewing yeasts (52). Because of their close phylogenetic relationship, CDFM21L.1, yHRVM108, and the *S. eubayanus* lager parent are more likely to share strain-specific genes, such as *lgAGT1* (64). From a search of Illumina sequencing reads available for CDFM21L.1 and yHRVM108, we were able to assemble two full-length genes with high sequence identity to *lgAGT1,* which we designated *tbAGT1* and *ncAGT1,* for Tibetan-*AGT1* and North Carolinian-*AGT1*, respectively (Fig 1).

Two single nucleotide polymorphisms (SNPs) separate *tbAGT1* and *lgAGT1.* One SNP results in a synonymous substitution and the other in a nonsynonymous substitution near the N-terminus of the protein outside of any predicted transmembrane domains (Fig 1B and C, S1 Fig). Analyses of the predicted effect of this substitution in *lgAGT1* (using STRUM and SIFT mutant protein prediction software (65,66)) suggest that it is unlikely to significantly impact protein structure or function (S1 Table). In contrast, *ncAGT1* has 95% nucleotide identity with *lgAGT1,* with nonsynonymous differences distributed throughout the sequence (Fig 1A-C). Despite the presence of *ncAGT1*, the yHRVM108 wild-type strain grows poorly on maltose and is unable to grow on maltotriose. Interestingly, and unlike all *MALT* genes found in the Patagonian type strain of *S. eubayanus*, overexpression of *ncAGT1* in yHRVM108 conferred growth on maltotriose, similar to the known maltotriose transporter gene *lgAGT1* (Table 1). These results suggest that insufficient *ncAGT1* gene expression, rather than protein function, is likely the main reason for the inability of yHRVM108 to grow on maltotriose.

Phylogenetic relationships among maltose and maltotriose transporters

To put the relationship between *S. eubayanus, S. cerevisiae,* and lager *MALT* genes into a phylogenetic perspective, a gene tree was constructed for these three groups of genes (Fig 2). Consistent with previous analyses of *MALT* genes in *Saccharomyces* (30), the *MALT* genes fell into 3 major clades. The *AGT1* genes formed their own group, significantly divergent from the other clades and was further split between the *AGT1* genes originating from *S. cerevisiae* and the *AGT1* genes originating

161 from *S. eubayanus*. *MPH* genes, which are native to *S. cerevisiae* but also present in some lager yeasts
162 (63,67), also formed their own clade. *MPH* genes are most often described as encoding maltose
163 transporters, but their ability to transport maltotriose is still uncertain (30,63,67–69). The final and
164 largest clade was made up of *MALT1-4* from *S. eubayanus, MALx1* genes from *S. cerevisiae,* and the
165 lager-specific gene *MTT1* (28,29,54). This clade was further subdivided into a group containing only *S.*
166 *eubayanus MALT* genes and their close lager homologs and another group consisting of *MALx1* genes,
167 *MTT1*, and *MALT3*. Within this clade, genes encoding maltotriose transporters were rare, represented
168 by only a single gene, *MTT1* (28,29). The phylogenetic distribution of maltotriose utilization suggests
169 that the ability to transport maltotriose may be a difficult function for genes within this clade to evolve.

170 **Indirect evolution of maltotriose consumption**

171 Since yHRVM108 contains a functional maltotriose transporter whose overexpression is
172 sufficient for growth on maltotriose (Table 1), we anticipated that it would be simple for yHRVM108 to
173 evolve the ability to utilize maltotriose under direct selection for this trait. Because yHRVM108 is unable
174 to grow on maltotriose, we passaged the strain in 2% maltotriose medium with a small amount of added
175 glucose to permit a limited number of cell divisions to allow for mutation and selection to occur. Over
176 the course of 100 passages under this selection regime, representing around 1,050 cell divisions across
177 three experimental replicates, no maltotriose-utilizing lineage of yHRVM108 arose.

178 While evolving yHRVM108 directly for maltotriose consumption was not successful, we were
179 initially surprised to find an alternative and indirect selection regime was effective at evolving
180 maltotriose utilization in this background. When we began adaptive evolution of yHRVM108 on
181 maltotriose, we also began selecting for increased growth of yHRVM108 on maltose to try and improve
182 this strain's sluggish growth on this carbon source. All three replicates of this experiment eventually
183 evolved the ability to grow rapidly on maltose. Interestingly, in addition to growing on maltose four
184 times more rapidly over two days (S2 Table), single-colony isolates from the first two replicates that

185 evolved rapid maltose utilization also gained the ability to utilize maltotriose (S3 Table), despite never
186 being exposed to maltotriose during the course of the adaptive evolution experiment. The fact that
187 maltotriose consumption independently evolved during selection for improved maltose utilization
188 multiple times, most likely through increased expression of *ncAGT1*, suggests that our maltotriose
189 selection regime itself may have played a role in restraining evolution.

190 Though we found the difficulty of evolving expression of a functional transporter surprising,
191 such a result is not unprecedented. In a long-term evolution experiment in *Escherichia coli*, a functioning
192 citrate transporter was present in the founding strain. Though expression of this gene would have been
193 highly favored in the citrate-rich experimental environment, it took thousands of generations, even after
194 the necessary potentiating mutations had appeared, before a gene amplification and rearrangement
195 event joined the citrate transporter gene to a new promoter, resulting in a novel expression pattern
196 (70). These results show how an organism's preexisting genetic architecture, interacting with the
197 selective environment, can facilitate or impede evolution along a particular path (71–74). Since maltose
198 is consumed to some extent by yHRVM108, loosened regulation of *ncAGT1* might have been selected for
199 alongside other *MAL* genes because *AGT1*-type transporters have a broad substrate range that includes
200 maltose (30). In contrast, evolving maltotriose utilization directly might have required specific and rare
201 changes to the *ncAGT1* locus itself. In retrospect, what appeared to be a simple request, to turn on the
202 *ncAGT1* gene in the presence of maltotriose as the sole carbon source, may in fact have been quite
203 difficult by simple mutations, whereas our indirect selection regime on maltose proved more effective.

204 **Evolution of maltotriose utilization through a chimeric transporter**

205 To determine how strains lacking any maltotriose transporters could evolve them, we also tried
206 to experimentally evolve maltotriose utilization in the *S. eubayanus* strains FM1318 (48) and in yHKS210
207 (53). A search of the available genome sequence reads for FM1318 (54,59) and yHKS210 (52) confirmed
208 that neither of these strains contain genes that are closely related to *AGT1*-like genes or other known

209 maltotriose transporters (28,29). Based on our analysis of the available whole-genome sequencing data,
210 these strains only contain the four *MALT* transporter genes previously identified in FM1318 (54), which
211 are unable to confer maltotriose utilization even when overexpressed (Table 1). Since, like yHRVM108,
212 neither of these strains could grow on maltotriose, a small amount of glucose was also added to the
213 medium to permit a limited number of cell divisions for mutation and selection. Over the course of 100
214 passages, representing approximately 2,100 cell divisions in total between the two strains and their
215 replicates, a single replicate, derived from strain yHKS210, evolved the ability to grow on maltotriose.
216 Two single-colony isolates (yHEB1505-6) from this replicate were isolated and confirmed to be able to
217 grow on maltotriose without added glucose (Fig 3A, S3 Table).
218 To determine the genetic architecture of maltotriose utilization in the replicate of yHKS210 that
219 evolved the ability to grow on maltotriose, we set up an F_1 backcross between the evolved maltotriose-
220 utilizing isolate yHEB1505 and the parent strain (yHKS210), producing strain yHEB1593, a putative
221 heterozygote capable of growth on maltotriose (Fig 3B and E). In a test of 15 fully viable F_2 tetrads,
222 maltotriose utilization segregated in a perfect 2:2 manner (Fig 3C). These results suggest that the ability
223 of the evolved strain to utilize maltotriose is conferred by a dominant mutation at a single genetic locus.
224 We performed bulk segregant analysis (75–77) using strains derived from the F_2 spores, dividing them
225 between those that could (MalTri+) and those that could not (MalTri-) utilize maltotriose (Fig 3C), with a
226 total of 30 strains in each category. Twelve 1-kb regions were identified as potentially containing fixed
227 differences between the MalTri+ and MalTri- strains. Of these regions, eight mapped to genes encoding
228 ribosomal proteins and most likely represent assembly artefacts due to the presence of many closely
229 related paralogs and/or their absence from the MalTri- de novo assembly that was used for
230 comparisons. Three other regions contained fixed differences between the MalTri+ and MalTri- groups
231 but had no clear relationship to carbon metabolism. The final 1-kb region mapped to the *MALT4* locus of
232 *S. eubayanus* genome (54,59). The coding sequence of *MALT4* from the MalTri+ group contained 52

SNPs relative to the *MALT4* allele found in yHKS210, all of which occurred within a single 230-bp region. Of these, 11 were predicted to lead to non-synonymous changes. Closer inspection revealed that the changes within the 230-bp region were the result of a recombination event between *MALT4* and *MALT3,* creating a chimeric gene (Fig 4), likely through ectopic gene conversion. We call this chimeric *MALT4* allele *MALT434* after the arrangement of sequences from its parent genes. The sequence of *MALT3* was not impacted by this mutational event.

To confirm that *MALT434* was the causative locus of maltotriose utilization, we performed a reciprocal hemizygosity test (78) in the heterozygous F_1 backcross strain (Fig 3D). Removal of *MALT434* eliminated the F_1 backcross strain's ability to utilize maltotriose (Fig 3E), demonstrating that *MALT434* is required for maltotriose utilization. Conversely, removing the parental, non-chimeric allele of *MALT4* in the heterozygous F_1 backcross strain had no impact on maltotriose utilization. Furthermore, overexpression of Malt434 in both the unevolved parent, yHKS210, and in the yHRVM108 background (Fig 5) supported growth on maltotriose, demonstrating that overexpression of Malt434 is sufficient to confer maltotriose utilization. These results strongly suggest that the mutant *MALT434* gene encodes a functional maltotriose transporter.

Potential structural impact of Malt434 chimerism

It was surprising that sequences from *MALT3* enabled *MALT4* to encode a maltotriose transporter because neither *MALT3* nor *MALT4* supported maltotriose utilization on their own (Table 1). Malt3 and Malt4 share about 80% amino acid identity overall and 85% amino acid identity in the chimeric region specifically (Fig 4B). Most residues in the chimeric region had high similarity between Malt3 and Malt4, as measured by Blosum62 similarity matrix (Fig 4C) (79), but there were a handful of low-similarity amino acids as well. To gain insight into what changes in protein structure may be driving the new functionality of Malt434, we used I-TASSER (80–82) to predict the protein structure of Malt3, Malt4, and Malt434. I-TASSER predicts a protein's structure based on its homology to proteins whose

structures have already been solved. Consistent with other studies on the structure of maltose transporters in *Saccharomyces* (27,83–85), I-TASSER predicted that Malt3, Malt4, and Malt434 were similar to members of the Major Facilitator Superfamily (MFS) of transporters, specifically the sugar porter family (85). Protein structure is predicted to be conserved between Malt3 and Malt4, including within the chimeric region, which encompasses one full transmembrane domain and parts of two other transmembrane domains (Fig 4D). Four maltose-binding sites were also predicted in the chimeric region. These same domains and predicted binding residues were predicted for Malt434 as well. Interestingly, I-TASSER predicted several of the alpha helices to be shorter in the chimera relative to the parent proteins: two alpha helices in the chimeric region and two towards the N-terminal end of the protein (Fig 4D, S1 Fig). The regions covered by these alpha helices were otherwise predicted to be conserved, out to phylogenetically distantly related Malt proteins lgAgt1 and scAgt1 (Fig 2, Fig 4D, S1 Fig). The predicted shortening of some alpha helices suggests that recombining the *MALT3* region into *MALT4* may have decreased the overall rigidity of the encoded chimeric protein, allowing it to accommodate bulkier substrates, such as maltotriose. Mutations that increase structural flexibility have been recognized in protein engineering as an important step in accommodating new substrates (86,87).

Besides increasing overall flexibility, the specific location of the chimeric region could have also played a role in supporting maltotriose transport. A previous study found that two residues were important for scAgt1's ability to transport maltotriose, while not affecting its ability to transport maltose (44). One of these residues lies within the chimeric region of Malt434, and the other is 10 amino acids downstream (Fig 4D, S1 Fig). Since the overall structure of maltose/maltotriose transporters is conserved (27,83–85), the area in and around the chimeric region of Malt434 may itself be important for substrate specificity.

Thus, the chimeric structure of Malt434 may have facilitated maltotriose transport in two ways. First, it may have increased the overall flexibility of the protein, allowing it to accommodate the larger

maltotriose molecule. Second, it could also have specifically altered an important substrate interface to facilitate a better interaction with maltotriose, possibly also by making this region more flexible. Testing these biophysical and structural models will require future experiments, such as solving the crystal structures for Malt3, Malt4, and Malt434 as complexes with maltose and/or maltotriose.

A non-modular chimeric path to novel substrate utilization

Most of the work on functional innovations by chimeric proteins has focused on the rearrangement of discrete functional units, with or without the benefit of intronic sequences (6,7,14,88–91). However, Malt434 does not fit easily into the framework of new protein creation by the reordering/exchanging of modules, even when considering smaller functional units such as a single alpha helix. While the chimeric region does completely move one alpha helix from Malt3 into the Malt4 background, the breakpoints of the gene conversion also result in two other alpha helices with some residues from the Malt4 parent and some from the Malt3 parent, creating chimeric alpha helices (Fig 4D, S1 Fig). In addition, while domains important for sugar specificity probably exist in Malt3 and Malt4 (44,84), with respect to maltotriose, the “sugar specificity” domain(s) between Malt3 and Malt4 do not seem to have different functions or specificities in their native backgrounds. In Malt3 and Malt4, there are no specific “maltotriose-transporting” domains to be swapped. Instead, the ability of the residues from Malt3 to facilitate maltotriose transport likely relies on their interaction with one or more residues in Malt4, not on their independent ability to interact with maltotriose.

Rather than the modular framework of novel protein formation, Malt434 exemplifies another framework for how recombination can lead to the evolution of novel functions. Theoretical and experimental work has demonstrated the important role that recombination between related proteins can play in facilitating the evolution of new functions (3,4,15,92). Indeed, protein engineering has utilized the technique of DNA shuffling since the mid-1990’s to recombine closely related coding sequences to efficiently generate proteins with novel or improved functions (19). More recently,

experimental work has begun to demonstrate the importance of recombination between closely related proteins in nature for the evolution of new functions (15,18,92). In this model, two duplicate proteins neutrally accumulate the multiple amino acid changes needed for a new function independently. All of the mutations that are needed for the new function are then brought together at once, en masse, through recombination. This molecular mechanism allows proteins to "tunnel" to new functions, bypassing potentially deleterious intermediates that would be encountered through a series of amino acid substitutions (3,4).

While *MALT3* and *MALT4* are not recent duplicates, they are distant paralogs (Fig 2). In addition, as members of the sugar porter subfamily of proteins, they share a highly conserved protein structure (27,83–85). The conservative nature of sugar porter family proteins means that recombination events like the one that formed Malt434, which do not fall between clear domains, probably have a relatively high likelihood of creating functional transporters (93), albeit ones of unpredictable specificity. In the case of Malt434, we do not yet know which specific amino acid interactions were important for the gain of maltotriose utilization in the chimera, let alone the function or history of the residues in the background of their native protein sequences. It may be that they represent neutral changes in their parental backgrounds, but they also could have been selected for other specificities. Nevertheless, the independent accumulation of these changes in a common ancestral protein background eventually allowed these sequences to recombine and create a novel function.

Conclusions

Our findings suggest that the evolution of maltotriose utilization by *Saccharomyces* yeasts is not a straightforward process. Even when a functioning maltotriose transporter is available in the parent genome, the regulatory changes necessary to support atypical expression may be difficult to evolve under certain experimental conditions. Conversely, when a maltotriose transporter is not already present, single point mutations are probably insufficient to switch or expand the specificity of available

Malt proteins. Recombination between paralogous proteins can rapidly do what a single point mutation cannot and, in a single rare mutational event, introduce the multiple residue changes needed to perform a new function. Our report on the evolution of a chimeric maltotriose transporter from parental proteins that could not transport maltotriose supports the role of recombination, beyond the simple swapping of functional protein domains and peptide motifs, in the formation of proteins with novel functions.

Materials and Methods

Strains

All strains discussed in this paper are listed in S4 Table. Briefly, FM1318 is a monosporic derivative of the type strain of *S. eubayanus*, which was isolated from Patagonia (48). yHRVM108 was isolated from Durham, North Carolina, and is closely related to the *S. eubayanus* strains that hybridized with *S. cerevisiae* to give rise to lager-brewing yeasts (52). yHKS210 was isolated from Sheboygan, Wisconsin, and is the result of admixture between populations A and B of *S. eubayanus*. yHKS210 is nearly homozygous due to selfing after the initial admixture event (53). Of these strains, FM1318 and yHKS210 grew well on maltose, but they did not grow on maltotriose. yHRVM108 grew sluggishly on maltose and did not grow on maltotriose. yHAB47 is a copy of Weihenstephan 34/70 (52), a representative of the Frohberg or Group II (94) lineage of lager-brewing hybrids (*S. cerevisiae* (2n) x *S. eubayanus* (2n) (58)). CDFM21L.1 is a strain of *S. eubayanus* isolated from Tibet (51) and is closely related to yHRVM108. Of known *S. eubayanus* strains, CDFM21L.1 is the most genetically similar to the *S. eubayanus* parents of lager-brewing hybrids (51,52).

Identification of *MALT* genes

Previously, we identified four genes with homology to genes encoding maltose transporters in *S. cerevisiae* and lager-brewing hybrids in the genome assembly of FM1318 published by Baker *et al.* 2015

(54). These genes were previously designated *MALT1-4.* Only a partial contig was available for *MALT4* in this assembly, but a BLAST (95) search of the Okuno *et al.* 2016 (59) assembly of the type strain of *S. eubayanus* (of which FM1318 is a monosporic derivative) allowed us to annotate the full-length sequence of *MALT4*. *MALT4* has 99.7% identity to *MALT2* at the nucleotide level and is predicted to have 100% identity at the amino acid level. The regions from 900 bp downstream of *MALT2* and *MALT4* and upstream to the ends of chromosomes V and XVI (regions of approximately 12 kb in the Okuno *et al.* 2016 (59) assembly), respectively, share 99.1% nucleotide identity. The 10 kb outside of this region only share 49.8% nucleotide identity. Thus, *MALT2* and *MALT4* are close paralogs that are likely related by a recent subtelomeric duplication and/or translocation event.

Reads for homologs of *AGT1* were retrieved using the functional *AGT1* sequence from lager yeast (*lgAGT1*) as the query sequence (55) in an SRA-BLAST search of the SRA databases of NCBI for yHRVM108 (SRR2586159) and CDFM21L.1 (SRR1507225). All reads identified in the BLAST searches were downloaded and assembled using the de novo assembler in Geneious v. 9.0.3 (http://www.geneious.com) (96). The homologs identified in yHRVM108 and CDFM21L.1 were designated *ncAGT1* (for North Carolinian *AGT1*) and *tbAGT1* (for Tibetan *AGT1*), respectively. The presence and sequence of *ncAGT1* in yHRVM108 was further verified by PCR amplification and Sanger sequencing (S5 Table). CDFM21L.1 was not available at the time of this work for further verification of the presence of *tbAGT1*.

Adaptive evolution

Design of the adaptive evolution experiments was based on Parreiras *et al.* 2014 (97). The highest available purities of carbon sources were used: 98% pure maltotriose, ≥99% pure maltose, and 100.0% pure glucose. Adaptive evolution was initiated by growing parent strains overnight in liquid YPD medium (1% yeast extract, 2% peptone, 2% glucose). One mL of maltotriose or maltose medium was inoculated with enough overnight culture to give an OD_{600} reading of ~0.1, as measured with an IMPLEN

OD600 DiluPhotometer™. Evolution on maltotriose was conducted in synthetic complete (SC) medium (0.17% yeast nitrogen base, 0.5% ammonium sulfate, 0.2% complete drop out mix) with 2% maltotriose and 0.1% glucose. The addition of 0.1% glucose ensured enough growth that mutations could occur and be selected for through the ensuing generations. Adaptive evolution of yHRVM108 on maltose was carried out in SC with 2% maltose. Because yHRVM108 grew so poorly on maltose alone, an additional 0.1% glucose was supplemented into its medium; after increased growth was observed around generation 110 for replicate A (from which strains yHEB1585-1587 were derived), around generation 80 for replicate B (from which strains yHEB1588-90 were derived), and around generation 155 for replicate C (from which strains yHEB1778-80 were derived), subsequent generations of yHRVM108 adaptive evolution on maltose for these replicates were conducted with 2% maltose only. Adaptive evolution experiments of each strain were carried out in triplicate. Samples were grown on a culture wheel at room temperature (22˚C) and diluted 1:10 into fresh media every 3-4 days. Samples of each evolution replicate were taken every other passage and placed into long-term storage by mixing 700uL of culture with 300uL of 50% glycerol in a cryotube and storing it at -80˚C. The numbers of doublings between passages were estimated from cell counts during the second and third passages. Evolution was carried out for a total of 100 passages. Strains that could not use the primary carbon source in the adaptive evolution medium underwent approximately one cell division per day on average.

Sporulation and backcrossing

To induce sporulation, strains were grown to saturation, washed twice, and then resuspended in 200µL liquid sporulation (spo) medium (1% potassium acetate, 0.5% zinc acetate). 30µL of this suspension was added to 1.5mL of spo medium and incubated on a culture wheel at room temperature. Cultures were checked for sporulation after 2-5 days. Tetrads were dissected using a Singer SporePlay. For backcrossing, tetrads of the strains to be crossed were dissected on a single YPD plate. A spore from one parent was placed in close proximity to a spore from the other parent, and they were observed over

400 several hours for mating and zygote formation. Transformations of the diploid F_1 backcross strain for
401 gene knockouts were carried out as described below in the section describing the construction of gene
402 expression plasmids.

403 **Construction of gene expression plasmids**

404 Genes encoding transporters of interest were cloned via gap repair into the *Not*I site of plasmid
405 pBM5155 (GenBank KT725394.1), which contains the complete machinery necessary for doxycycline-
406 based induction of genes cloned into this site (98). Transformation was carried out using standard
407 lithium acetate transformation (99) with modifications to optimize transformation in *S. eubayanus*.
408 Specifically, transformation reactions were heat-shocked at 34°C. After 55 minutes, 100% ethanol was
409 added to 10% total volume, and the reactions heat shocked for another 5 minutes before they were
410 allowed to recover overnight and plated to selective media the next day. When necessary, plasmids
411 were recovered and amplified in *E. coli* for transformation into multiple strains. The sequences of genes
412 encoding transporters cloned into pBM5155 were verified by Sanger sequencing. *S. eubayanus MALT1*,
413 *MALT3*, and *MALT4* were amplified from FM1318, *lgAGT* was amplified from yHAB47, and *ncAGT1* was
414 amplified from yHRVM108. Primers used for plasmid construction and sequence verification are listed in
415 S5 Table.

416 **Growth assays**

417 Growth was measured in liquid media in 96-well plates using OD_{600} measurements on a
418 FLUOstar Omega® microplate reader. Strains were first grown to saturation in liquid YPD medium, then
419 washed twice and diluted in SC without added carbon to OD_{600} = 1.9 +/- 0.05 to ensure that all cultures
420 had approximately the same starting concentration. 15µL of each diluted culture was added to 235µL of
421 the test medium. Three technical replicates, randomly distributed on a 96-well plate to control for
422 position effects, were carried out for each strain. Single-colony isolates of yHKS210 evolved on
423 maltotriose and single-colony isolates of yHRVM108 evolved on maltose were tested in SC medium + 2%

maltotriose. Single-colony isolates of yHRVM108 evolved on maltose were also tested on SC medium + 2% maltose. Strains carrying *MALT* genes expressed on an inducible plasmid were tested in SC medium + 2% maltotriose and 5 ng/mL doxycycline to induce plasmid gene expression. To control for growth from the small amount of non-maltotriose sugar in 98% pure maltotriose, the parent strains of yHRVM108 and yHKS210 were also tested in SC medium + 0.04% glucose, reflecting the approximate amount of other carbon sources expected in SC medium + 2% maltotriose.

Bulk segregant analysis

60 spores from 15 fully viable tetrads of strain yHEB1593 (F_1 of yHKS210 x yHEB1505) were dissected and individually screened for their ability to grow in SC + 2% maltotriose. F_2 segregants that could grow on maltotriose were classified as MalTri+, and those that could not were classified as MalTri-. Each F_2 segregant was then individually grown to saturation in liquid YPD. The saturated cultures were spun down, the supernatant removed, and enough cells resuspended in liquid SC medium to give an OD_{600} measurement of between 1.9 and 1.95, as measured with an IMPLEN OD600 DiluPhotometer™. Strains were pooled based on their ability to grow on maltotriose, forming a MalTri+ pool and a MalTri- pool. To pool, 1mL of each strain dilution was added to the appropriate pool of cells and vortexed to mix. Phenol-chloroform extraction and ethanol precipitation was used to isolate gDNA from the segregant pools. The gDNA was sonicated and ligated to Illumina TruSeq-style dual adapters and index sequencing primers using the NEBNext® DNA Library Prep Master Mix Set for Illumina® kit following the manufacturer's instructions. The paired-end libraries were sequenced on an Illumina MiSeq instrument, conducting a 2 x 250bp run.

Analysis of bulk segregant sequencing reads

To identify fixed differences between the meiotic segregant pools, de novo assemblies were made for the MalTri- group of segregants using the meta-assembler iWGS with default settings (100). The final genome assembly of the MalTri- pool was made by DISCOVAR (101) in iWGS. This assembly was

448 used for reference-based genome assembly and variant calling using reads from the MalTri+ pool
449 following the protocol described in Peris and Langdon *et al.* 2016 (52). Assemblies of the putative
450 chimeric maltotriose transporter were retrieved from the MalTri+ pool of reads using the program
451 HybPiper (102). Briefly, HybPiper uses a BLAST search of read sequences to find reads that map to a
452 query sequence; it then uses the programs Exonerate (103) and SPAdes (104) to assemble the reads into
453 contigs. The sequence and genomic location of the chimeric transporter were further verified by PCR
454 amplification and Sanger sequencing (S5 Table), as was the sequence of *MALT4* from yHKS210.

455 **Phylogenetic analyses and computational predictions of protein structures and functions**

456 Multiple sequence alignments between the proteins encoded by the *MALT* genes were carried
457 out using MUSCLE (105), as implemented in Geneious v.9.0.3 (96) (http://www.geneious.com).
458 Phylogenetic relationships were determined using codon alignments. Codon alignments were made
459 using PAL2NAL (Suyama, Torrents, & Bork, 2006; http://www.bork.embl.de/pal2nal/) to convert the
460 MUSCLE alignments of amino acid sequences to nucleotide alignments. A phylogenetic tree of nineteen
461 *MALT* genes from *S. eubayanus* and *S. cerevisiae* and lager-brewing yeasts was constructed as described
462 in Baker *et al.* 2015 (54) using MEGA v.6. Most genes used in the phylogenetic analysis were retrieved as
463 previously described in Baker *et al.* 2015 (54) as follows: *MAL21, MAL31*, and *MAL61* from *S. cerevisiae*;
464 *MALT1* and *MALT3* from *S. eubayanus*; *MALT1, MALT2,* and *MPH* from lager-brewing yeast; *MPH2 and*
465 *MPH3* from *S. cerevisiae*; *AGT1* (*MAL11* in Baker *et al.* 2015 (54)) from *S. cerevisiae*; *scAGT1*
466 (*WeihenMAL11*-CB in Baker *et al*. 2015 (54)); and *lgAGT1* (*WeihenMAL11*-CA in Baker *et al*. 2015 (54))
467 Sequences for *MALT2* and *MALT4* were retrieved from the genome assembly of CBS 12357^{T} from Okuno
468 *et al.* 2016 (59). *MAL11* was retrieved from the genome assembly of *S. cerevisiae* strain YJM456 (107).
469 Sequences for *tbAGT1* and *ncAGT1* were retrieved as described above. *MAL11* and *AGT1* both encode α-
470 glucoside transporters located at the *MAL1* locus in *S. cerevisiae* and, as such, are considered alleles of
471 each other (27,108). Their shared genomic location notwithstanding, *MAL11* and *AGT1* are not

phylogenetically closely related, with *MAL11* clustering with other *MALx1* type transporters (Fig 2). In addition, while *AGT1* can support maltotriose transport, *MAL11,* like other known *MALx1* genes, cannot (27,30). Despite their dissimilarity, *AGT1* is recorded in the *Saccharomyces* Genome Database (yeastgenome.org) as *MAL11* since the reference strain carries the *AGT1* allele at the *MAL1* locus (60,63). For this reason, *MAL11* is often used to refer to *AGT1* (30,32,54). For clarity, here we use *MAL11* to only refer to the *MALx1*-like allele and *AGT1* to refer to the distinct maltotriose-transporting allele.

Protein structure predictions for *MALT3*, *MALT4*, *lgAGT1*, and *scAGT1* were carried out using the I-TASSER server, and the structure prediction of *MALT434* was carried out using the command line version of I-TASSER (80–82) (https://zhanglab.ccmb.med.umich.edu/I-TASSER/, accessed between 2-7-2018 and 2-28-2018). The potential impact of the single residue difference between *lgAGT1* and *tbAGT1* was analyzed by two different methods. Prediction of the change in free energy ($\Delta\Delta G$) was carried out using the STRUM server (https://zhanglab.ccmb.med.umich.edu/STRUM/, accessed 3-21-18) (65). A $\Delta\Delta G$ score of < +/- 0.5 was considered to be unlikely to affect function (109). Homology-based predictions were made using SIFT at http://sift.jcvi.org/ (accessed 3-30-18) (66,110–113). The SIFT Related Sequences analysis was done using the amino acid sequences of *MALT* genes in the phylogenetic analysis above. Several SIFT analyses were also carried out using the SIFT Sequence analysis program. This analysis operates using the same principle as the SIFT Related Sequences analysis, but rather than being supplied by the user, homologous sequences were provided by a PSI-BLAST search of the indicated protein database. The SIFT Sequence analyses were carried out using default settings and the following databases available on http://sift.jcvi.org/ (accessed 3-30-18): NCBI nonredundant 2011 Mar, UniRef90 2011 Apr, UniProt-SwissProt 57.15 2011 Apr.

Acknowledgments

We thank Dana A. Opulente, Quinn K. Langdon, Ryan V. Moriarty, and Kelly V. Buh for their assistance maintaining the experimental evolution lines described in this manuscript.

Conflicts of interest

I have read the journal's policy and the authors of this manuscript have the following competing interests: EB and CTH, together with the Wisconsin Alumni Research Foundation, have filed a provisional patent application entitled, "POLYPEPTIDE AND YEAST CELL COMPOSITIONS AND METHODS OF USING THE SAME." All strains and constructs are freely available for non-commercial research.

Financial Disclosure

This work was supported by the USDA National Institute of Food and Agriculture, Hatch project 1003258, the National Science Foundation (grant numbers DEB-1253634); and funded in part by the DOE Great Lakes Bioenergy Research Center (DOE BER Office of Science DE-SC0018409 and DE-FC02-07ER64494). EB is supported by a Louis and Elsa Thomsen Wisconsin Distinguished Graduate Fellowship. CTH is a Pew Scholar in the Biomedical Sciences and a Vilas Faculty Early Career Investigator, supported by the Pew Charitable Trusts and the Vilas Trust Estate. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Long M, Betrán E, Thornton K, Wang W. The origin of new genes: glimpses from the young and old. Nat Rev Genet. 2003;4(11):865–75. pmid:14634634
2. Ohno S. Evolution by Gene Duplication. Berlin, Heidelberg: Springer Berlin Heidelberg; 1970.
3. Bittihn P, Tsimring LS. Gene Conversion Facilitates Adaptive Evolution on Rugged Fitness Landscapes. Genetics. 2017;207(4):1577–89. pmid:28978673
4. Cui Y, Wong WH, Bornberg-Bauer E, Chan HS. Recombinatoric exploration of novel folded structures: a heteropolymer-based model of protein evolutionary landscapes. Proc Natl Acad Sci U S A. 2002;99(2):809–14. pmid:11805332
5. Poelwijk FJ, Kiviet DJ, Weinreich DM, Tans SJ. Empirical fitness landscapes reveal accessible evolutionary paths. Nature. 2007;445(7126):383–6.
6. Patthy L. Genome evolution and the evolution of exon-shuffling — a review. Gene. 1999;238(1):103–14. pmid:10570989
7. Patthy L. Modular Assembly of Genes and the Evolution of New Functions. Genetica. 2003;118(2–

3):217–31. pmid:12868611

8. Henikoff S, Greene EA, Pietrokovski S, Bork P, Attwood TK, Hood L. Gene families: the taxonomy of protein paralogs and chimeras. Science. 1997;278(5338):609–14. pmid:9381171

9. Leffler EM, Band G, Busby GBJ, Kivinen K, Le QS, Clarke GM, et al. Resistance to malaria through structural variation of red blood cell invasion receptors. Science. 2017;356(6343):eaam6393. pmid:28522690

10. Mitelman F, Johansson B, Mertens F. The impact of translocations and gene fusions on cancer causation. Nat Rev Cancer. 2007;7(4):233–45. pmid:17361217

11. Rippey C, Walsh T, Gulsuner S, Brodsky M, Nord AS, Gasperini M, et al. Formation of Chimeric Genes by Copy-Number Variation as a Mutational Mechanism in Schizophrenia. Am J Hum Genet. 2013;93(4):697–710. pmid:24094746

12. Malfavon-Borja R, Wu LI, Emerman M, Malik HS. Birth, decay, and reconstruction of an ancient TRIMCyp gene fusion in primate genomes. Proc Natl Acad Sci U S A. 2013;110(7):E583-92. pmid:23319649

13. Wang W, Zhang J, Alvarez C, Llopart A, Long M. The Origin of the *Jingwei* Gene and the Complex Modular Structure of Its Parental Gene, Yellow Emperor, in *Drosophila melanogaster*. Mol Biol Evol. 2000;17(9):1294–301.

14. De Chateau M, Bjorck L, Smith JM. Identification of interdomain sequences promoting the intronless evolution of a bacterial protein family. Proc Natl Acad Sci. 1996;93(16):8490–5. pmid:8710897

15. Rogers RL, Hartl DL. Chimeric Genes as a Source of Rapid Evolution in *Drosophila melanogaster*. Mol Biol Evol. 2012;29(2):517–29. pmid:21771717

16. Bogarad LD, Deem MW. A hierarchical approach to protein molecular evolution. Proc Natl Acad Sci U S A. 1999;96(6):2591–5. pmid:10077554

17. Voigt CA, Martinez C, Wang Z-G, Mayo SL, Arnold FH. Protein building blocks preserved by recombination. Nat Struct Biol. 2002;9(7):553–8. pmid:12042875

18. Thomas JH. Concerted evolution of two novel protein families in *Caenorhabditis* species. Genetics. 2006;172(4):2269–81. pmid:16415360

19. Giver L, Arnold FH. Combinatorial protein design by in vitro recombination. Curr Opin Chem Biol. 1998;2(3):335–8. pmid:9691067

20. Cole MF, Gaucher EA. Utilizing natural diversity to evolve protein function: applications towards thermostability. Curr Opin Chem Biol. 2011;15(3):399–406.

21. Minshull J, Willem Stemmer PC. Protein evolution by molecular breeding. Curr Opin Chem Biol. 1999;3(3):284–90. pmid:10359711

22. Nijland JG, Shin HY, de Waal PP, Klaassen P, Driessen AJM. Increased xylose affinity of Hxt2 through gene shuffling of hexose transporters in *Saccharomyces cerevisiae*. J Appl Microbiol. 2018 Feb;124(2):503–10. pmid:29240974

23. Eßlinger HM. Fermentation, Maturation and Storage. In: Eßlinger HM, editor. Handbook of brewing processes, technology, markets. Weinheim (Germany): Wiley-VCH; 2009. p. 209.

24. Briggs DE, Brookes PA, Stevens R, Boulton CA. Metabolism of wort by yeast. In: Brewing: science and practice. Abington Hall, Abington Cambridge CB1 6AH, England: Woodhead Publishing Limited; 2004. p. 413.

25. Briggs D.E., Brookes P.A., Stevens R. BCA. The science of mashing. In: Brewing: science and practice. Abington Hall, Abington Cambridge CB1 6AH, England: Woodhead Publishing Limited; 2004. p. 108, 109, 123,125.

26. Meussdoerffer F, Zarnkow M. Starchy Raw Materials. In: Eßlinger HM, editor. Handbook of brewing processes, technology, markets. Weinheim (Germany): Wiley-VCH; 2009. p. 58.

27. Han E, Cotty F, Sottas C, Jiang H, Michels CA. Characterization of *AGT1* encoding a general

α-glucoside transporter from *Saccharomyces*. Mol Microbiol. 1995;17(6):1093–107. pmid:8594329

28. Dietvorst J, Londesborough J, Steensma HY. Maltotriose utilization in lager yeast strains: *MTT1* encodes a maltotriose transporter. Yeast. 2005;22(10):775–88. pmid:16088872

29. Salema-Oom M, Valadão Pinto V, Gonçalves P, Spencer-Martins I. Maltotriose utilization by industrial *Saccharomyces* strains: characterization of a new member of the alpha-glucoside transporter family. Appl Environ Microbiol. 2005;71(9):5044–9. pmid:16151085

30. Brown CA, Murray AW, Verstrepen KJ. Rapid Expanion and Functional Divergence of Subtelomeric Gene Families in Yeasts. Curr Biol. 2010;20(10):895–903. pmid:20471265

31. Horák J. Regulations of sugar transporters: insights from yeast. Curr Genet. 2013;59(1–2):1–31. pmid:23455612

32. Brickwedde A, van den Broek M, Geertman J-MA, Magalhães F, Kuijpers NGA, Gibson B, et al. Evolutionary Engineering in Chemostat Cultures for Improved Maltotriose Fermentation Kinetics in *Saccharomyces pastorianus* Lager Brewing Yeast. Front Microbiol. 2017;8:1690. pmid:28943864

33. Jansen MLA, Daran-Lapujade P, de Winde JH, Piper MDW, Pronk JT. Prolonged maltose-limited cultivation of *Saccharomyces cerevisiae* selects for cells with improved maltose affinity and hypersensitivity. Appl Environ Microbiol. 2004;70(4):1956–63. pmid:15066785

34. Krogerus K, Seppänen-Laakso T, Castillo S, Gibson B. Inheritance of brewing-relevant phenotypes in constructed *Saccharomyces cerevisiae* × *Saccharomyces eubayanus* hybrids. Microb Cell Fact. 2017;16(1):66. pmid:28431563

35. Krogerus K, Magalhães F, Vidgren V, Gibson B. New lager yeast strains generated by interspecific hybridization. J Ind Microbiol Biotechnol. 2015;42(5):769–78. pmid:25682107

36. Hebly M, Brickwedde A, Bolat I, Driessen MRM, de Hulster EAF, van den Broek M, et al. *S*.

cerevisiae × *S. eubayanus* interspecific hybrid, the best of both worlds and beyond. FEMS Yeast Res. 2015;15(3):fov005. pmid:25743788

37. Krogerus K, Arvas M, De Chiara M, Magalhães F, Mattinen L, Oja M, et al. Ploidy influences the functional attributes of de novo lager yeast hybrids. Appl Microbiol Biotechnol. 2016;100(16):7203–22. pmid:27183995

38. Mertens S, Steensels J, Saels V, De Rouck G, Aerts G, Verstrepen KJ. A large set of newly created interspecific *Saccharomyces* hybrids increases aromatic diversity in lager beers. Appl Environ Microbiol. 2015;81(23):8202–14. pmid:26407881

39. Nikulin J, Krogerus K, Gibson B. Alternative *Saccharomyces* interspecies hybrid combinations and their potential for low-temperature wort fermentation. Yeast . 2018;35(1):113–27. pmid:28755430

40. Stewart GG. The genetic manipulation of industrial yeast strains. Can J Microbiol. 1981;27(10):973–90.

41. Bilinski CA, Casey GP. Developments in sporulation and breeding of brewer's yeast. Yeast. 1989;5(6):429–38.

42. Mukai N, Nishimori C, Fujishige W, Mizuno A, Takahashi T, Sato K. Beer Brewing Using a Fusant between a Sake Yeast and a Brewer's Yeast. J Biosci Bioeng. 2001;91(5):482–6. pmid:16233026

43. Yamakawa S, Yamada R, Tanaka T, Ogino C, Kondo A. Repeated batch fermentation from raw starch using a maltose transporter and amylase expressing diploid yeast strain. Appl Microbiol Biotechnol. 2010;87(1):109–15. pmid:20180115

44. Smit A, Moses SG, Pretorius IS, Cordero Otero RR. The Thr505 and Ser557 residues of the AGT1-encoded alpha-glucoside transporter are critical for maltotriose transport in *Saccharomyces cerevisiae*. J Appl Microbiol. 2008;104(4):1103–11. pmid:18179544

45. Smit A, Dissertation. Maltotriose Transport in Yeast. Stellenbosch University; 2007. Avialible

from:
http://scholar.sun.ac.za/bitstream/handle/10019.1/21760/Smit_Maltotriose_1007.pdf?sequence=1&isAllowed=y

46. Krogerus K, Magalhães F, Vidgren V, Gibson B. Novel brewing yeast hybrids: creation and application. Appl Microbiol Biotechnol. 2017;101(1):65–78. pmid:27885413

47. Hittinger CT, Steele JL, Ryder DS. Diverse yeasts for diverse fermented beverages and foods. Curr Opin Biotechnol. 2018;49:199–206. pmid:29102814

48. Libkind D, Hittinger CT, Valério E, Gonçalves C, Dover J, Johnston M, et al. Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. Proc Natl Acad Sci U S A. 2011;108(35):14539–44. pmid:21873232

49. Gibson BR, Storgårds E, Krogerus K, Vidgren V. Comparative physiology and fermentation performance of Saaz and Frohberg lager yeast strains and the parental species *Saccharomyces eubayanus*. Yeast. 2013;30(7):255–66. pmid:23695993

50. Gibson B, Geertman J-MA, Hittinger CT, Krogerus K, Libkind D, Louis EJ, et al. New yeasts—new brews: modern approaches to brewing yeast design and development. FEMS Yeast Res. 2017;17(4). pmid:28582493

51. Bing J, Han P-J, Liu W-Q, Wang Q-M, Bai F-Y. Evidence for a Far East Asian origin of lager beer yeast. Curr Biol. 2014;24(10):R380-1. pmid:24845661

52. Peris D, Langdon QK, Moriarty R V., Sylvester K, Bontrager M, Charron G, et al. Complex Ancestries of Lager-Brewing Hybrids Were Shaped by Standing Variation in the Wild Yeast *Saccharomyces eubayanus*. PLOS Genet. 2016;12(7):e1006155. pmid:27385107

53. Peris D, Sylvester K, Libkind D, Gonçalves P, Sampaio JP, Alexander WG, et al. Population structure and reticulate evolution of *Saccharomyces eubayanus* and its lager-brewing hybrids. Mol Ecol. 2014;23(8):2031–45. pmid:24612382

643 54. Baker E, Wang B, Bellora N, Peris D, Hulfachor AB, Koshalek JA, et al. The Genome Sequence of
644 *Saccharomyces eubayanus* and the Domestication of Lager-Brewing Yeasts. Mol Biol Evol.
645 2015;32(11):2818–31. pmid:26269586
646 55. Nakao Y, Kanamori T, Itoh T, Kodama Y, Rainieri S, Nakamura N, et al. Genome sequence of the
647 lager brewing yeast, an interspecies hybrid. DNA Res. 2009;16(2):115–29. pmid:19261625
648 56. Vidgren V, Multanen J-P, Ruohonen L, Londesborough J. The temperature dependence of
649 maltose transport in ale and lager strains of brewer's yeast. FEMS Yeast Res. 2010;10(4):402–11.
650 pmid:20402791
651 57. Cousseau FEM, Alves SL, Trichez D, Stambuk BU. Characterization of maltotriose transporters
652 from the *Saccharomyces eubayanus* subgenome of the hybrid *Saccharomyces pastorianus* lager
653 brewing yeast strain Weihenstephan 34/70. Lett Appl Microbiol. 2013;56(1):21–9.
654 pmid:23061413
655 58. Nguyen H-V, Boekhout T. Characterization of *Saccharomyces uvarum* (Beijerinck, 1898) and
656 related hybrids: assessment of molecular markers that predict the parent and hybrid genomes
657 and a proposal to name yeast hybrids. FEMS Yeast Res. 2017;17(2).
658 59. Okuno M, Kajitani R, Ryusui R, Morimoto H, Kodama Y, Itoh T. Next-generation sequencing
659 analysis of lager brewing yeast strains reveals the evolutionary history of interspecies
660 hybridization. DNA Res. 2016;23(1):67–80. pmid:26732986
661 60. Vidgren V, Huuskonen A, Virtanen H, Ruohonen L, Londesborough J. Improved fermentation
662 performance of a lager yeast after repair of its *AGT1* maltose and maltotriose transporter genes.
663 Appl Environ Microbiol. 2009;75(8):2333–45. pmid:19181838
664 61. Vidgren V, Londesborough J. Characterization of the *Saccharomyces bayanus*-type *AGT1*
665 transporter of lager yeast. J Inst Brew. 2012;118(2):148–51.
666 62. Day RE, Rogers PJ, Dawes IW, Higgins VJ. Molecular analysis of maltotriose transport and

utilization by *Saccharomyces cerevisiae*. Appl Environ Microbiol. 2002;68(11):5326–35. pmid:12406721

63. Vidgren V, Ruohonen L, Londesborough J. Characterization and functional analysis of the *MAL* and *MPH* Loci for maltose utilization in some ale and lager yeast strains. Appl Environ Microbiol. 2005;71(12):7846–57. pmid:16332759

64. Bergström A, Simpson JT, Salinas F, Barré B, Parts L, Zia A, et al. A High-Definition View of Functional Genetic Variation from Natural Yeast Genomes. Mol Biol Evol. 2014;31(4):872–88.

65. Quan L, Lv Q, Zhang Y. STRUM: structure-based prediction of protein stability changes upon single-point mutation. Bioinformatics. 2016;32(19):2936–46. pmid:27318206

66. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. Genome Res. 2001;11(5):863–74. pmid:11337480

67. Day RE, Higgins VJ, Rogers PJ, Dawes IW. Characterization of the putative maltose transporters encoded by YDL247w and YJR160c. Yeast. 2002;19(12):1015–27. pmid:12210897

68. Stambuk BU, Alves SL, Hollatz C, Zastrow CR. Improvement of maltotriose fermentation by *Saccharomyces cerevisiae*. Lett Appl Microbiol. 2006;43(4):370–6. pmid: 16965366

69. Alves SL, Herberts RA, Hollatz C, Trichez D, Miletti LC, de Araujo PS, et al. Molecular analysis of maltotriose active transport and fermentation by *Saccharomyces cerevisiae* reveals a determinant role for the *AGT1* permease. Appl Environ Microbiol. 2008;74(5):1494–501. pmid:18203856

70. Blount ZD, Barrick JE, Davidson CJ, Lenski RE. Genomic analysis of a key innovation in an experimental *Escherichia coli* population. Nature. 2012;489(7417):513–8.

71. Fedoroff N V. Presidential address. Transposable elements, epigenetics, and genome evolution. Science. 2012;338(6108):758–67. pmid:23145453

72. Gresham D, Desai MM, Tucker CM, Jenq HT, Pai DA, Ward A, et al. The Repertoire and Dynamics

691 of Evolutionary Adaptations to Controlled Nutrient-Limited Environments in Yeast. Snyder M,
692 editor. PLoS Genet. 2008;4(12):e1000303. pmid: 19079573
693 73. Dunham MJ, Badrane H, Ferea T, Adams J, Brown PO, Rosenzweig F, et al. Characteristic genome
694 rearrangements in experimental evolution of *Saccharomyces cerevisiae*. Proc Natl Acad Sci U S A.
695 2002;99(25):16144–9. pmid:12446845
696 74. Liti G, Louis EJ. YEAST EVOLUTION AND COMPARATIVE GENOMICS. Annu Rev Microbiol.
697 2005;59(1):135–53. pmid: 15877535
698 75. Ehrenreich IM, Torabi N, Jia Y, Kent J, Martis S, Shapiro JA, et al. Dissection of genetically complex
699 traits with extremely large pools of yeast segregants. Nature. 2010;464(7291):1039–42.
700 pmid:20393561
701 76. Brauer MJ, Christianson CM, Pai DA, Dunham MJ. Mapping novel traits by array-assisted bulk
702 segregant analysis in *Saccharomyces cerevisiae*. Genetics. 2006;173(3):1813–6. pmid:16624899
703 77. Segrè A V, Murray AW, Leu J-Y. High-Resolution Mutation Mapping Reveals Parallel Experimental
704 Evolution in Yeast. PLoS Biol. 2006;4(8):e256. pmid:16856782
705 78. Stern DL. Identification of loci that cause phenotypic variation in diverse species with the
706 reciprocal hemizygosity test. Trends Genet. 2014;30(12):547–54. pmid:25278102
707 79. Henikoff S, Henikoff JG. Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci
708 U S A. 1992;89(22):10915–9. pmid:1438297
709 80. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and
710 function prediction. Nat Protoc. 2010;5(4):725–38. pmid:20360767
711 81. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function
712 prediction. Nat Methods. 2015;12(1):7–8. pmid:25549265
713 82. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 2008;9:40.
714 pmid:18215316

715 83. Cheng Q, Michels CA. The maltose permease encoded by the *MAL61* gene of *Saccharomyces*
716 *cerevisiae* exhibits both sequence and structural homology to other sugar transporters. Genetics.
717 1989;123(3):477–84. pmid:2689282
718 84. Barrett MP, Walmsleyt AR, Gould GW. Structure and function of facultative sugar transporters.
719 Curr Opin Cell Biol. 1999;11(4):496–502.
720 85. Yan N. Structural Biology of the Major Facilitator Superfamily Transporters. Annu Rev Biophys.
721 2015;44:257–83. pmid:26098515
722 86. Khersonsky Olga, Tawfik DS. Enzyme Promiscuity: A Mechanistic and Evolutionary Perspective.
723 Annu Rev Biochem. 2010;79(1):471–505. pmid:20235827
724 87. Mannige R V. Dynamic New World: Refining Our View of Protein Structure, Function and
725 Evolution. Proteomes. 2014;2(1):128–53. pmid:28250374
726 88. Bashton M, Chothia C. The Generation of New Protein Functions by the Combination of Domains.
727 Structure. 2007;15(1):85–99. pmid:17223535
728 89. Vogel C, Bashton M, Kerrison ND, Chothia C, Teichmann SA. Structure, function and evolution of
729 multidomain proteins. Curr Opin Struct Biol. 2004;14(2):208–16. pmid:15093836
730 90. Furuta Y, Kobayashi I. Movement of DNA sequence recognition domains between non-
731 orthologous proteins. Nucleic Acids Res. 2012;40(18):9218–32. pmid:22821560
732 91. Doolittle RF. The multiplicity of domains in proteins. Annu Rev Biochem. 1995;64:287–314.
733 pmid:7574483
734 92. Mody A, Weiner J, Ramanathan S. Modularity of MAP kinases allows deformation of their
735 signalling pathways. Nat Cell Biol. 2009;11(4):484–91. pmid:19295513
736 93. Drummond DA, Silberg JJ, Meyer MM, Wilke CO, Arnold FH. On the conservative nature of
737 intragenic recombination. Proc Natl Acad Sci U S A. 2005;102(15):5380–5. pmid:15809422
738 94. Magalhães F, Vidgren V, Ruohonen L, Gibson B. Maltose and maltotriose utilisation by group I

739 strains of the hybrid lager yeast *Saccharomyces pastorianus*. FEMS Yeast Res. 2016;16(5).
740 pmid:27364826
741 95. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-
742 BLAST: a new generation of protein database search programs. Nucleic Acids Res.
743 1997;25(17):3389–402. pmid:9254694
744 96. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: An
745 integrated and extendable desktop software platform for the organization and analysis of
746 sequence data. Bioinformatics. 2012;28(12):1647–9. pmid:22543367
747 97. Parreiras LS, Breuer RJ, Avanasi Narasimhan R, Higbee AJ, La Reau A, Tremaine M, et al.
748 Engineering and Two-Stage Evolution of a Lignocellulosic Hydrolysate-Tolerant *Saccharomyces*
749 *cerevisiae* Strain for Anaerobic Fermentation of Xylose from AFEX Pretreated Corn Stover. PLoS
750 One. 2014;9(9):e107499.
751 98. Alexander WG, Peris D, Pfannenstiel BT, Opulente DA, Kuang M, Hittinger CT. Efficient
752 engineering of marker-free synthetic allotetraploids of *Saccharomyces*. Fungal Genet Biol.
753 2016;89:10–7. pmid:26555931
754 99. Gietz DR, Woods RA. Transformation of yeast by lithium acetate/single-stranded carrier
755 DNA/polyethylene glycol method. Methods Enzymol. 2002;350:87–96. pmid:12073338
756 100. Zhou X, Peris D, Kominek J, Kurtzman CP, Hittinger CT, Rokas A. in silico Whole Genome
757 Sequencer and Analyzer (iWGS): A Computational Pipeline to Guide the Design and Analysis of de
758 novo Genome Sequencing Studies. G3. 2016;6(11):3655–62. pmid:27638685
759 101. Weisenfeld NI, Yin S, Sharpe T, Lau B, Hegarty R, Holmes L, et al. Comprehensive variation
760 discovery in single human genomes. Nat Genet. 2014;46(12):1350–5. pmid:25326702
761 102. Johnson MG, Gardner EM, Liu Y, Medina R, Goffinet B, Shaw AJ, et al. HybPiper: Extracting coding
762 sequence and introns for phylogenetics from high-throughput sequencing reads using target

763 enrichment. Appl Plant Sci. 2016;4(7). pmid:27437175

764 103. Slater G, Birney E. Automated generation of heuristics for biological sequence comparison. BMC
765 Bioinformatics. 2005;6:31. pmid:15713233

766 104. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome
767 assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19(5):455–
768 77. pmid:22506599

769 105. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic
770 Acids Res. 2004;32(5):1792–7. pmid:15034147

771 106. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into
772 the corresponding codon alignments. Nucleic Acids Res. 2006;34(Web Server):W609–12.
773 pmid:16845082

774 107. Strope PK, Skelly DA, Kozmin SG, Mahadevan G, Stone EA, Magwene PM, et al. The 100-genomes
775 strains, an *S. cerevisiae* resource that illuminates its natural phenotypic and genotypic variation
776 and emergence as an opportunistic pathogen. Genome Res. 2015;25(5):762–74. pmid:25840857

777 108. Charron MJ, Michels CA. The naturally occurring alleles of *MAL1* in *Saccharomyces* species
778 evolved by various mutagenic processes including chromosomal rearrangement. Genetics.
779 1988;120(1):83–93. pmid:2851483

780 109. Bromberg Y, Rost B. Correlating protein function and stability through the analysis of single
781 amino acid substitutions. BMC Bioinformatics. 2009;10(Suppl 8). pmid:19758472

782 110. Ng PC, Henikoff S. Accounting for human polymorphisms predicted to affect protein function.
783 Genome Res. 2002;12(3):436–46. pmid:11875032

784 111. Ng PC, Henikoff S. SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids
785 Res. 2003;31(13):3812–4. pmid:12824425

786 112. Ng PC, Henikoff S. Predicting the Effects of Amino Acid Substitutions on Protein Function. Annu

Rev Genomics Hum Genet. 2006;7:61–80. pmid:16824020

113. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2009;4(8):1073–81. pmid:19561590

Figure Legends

Fig 1. Alignment of *AGT1*-like genes. A) Tables highlighting the nucleotide (nuc) and amino acid (aa) percent identities between members of the *AGT1* family. Darker colors indicate greater sequence similarity. B) Multiple sequence alignment between nucleotide sequences of *tbAGT1, lgAGT1*, and *ncAGT1*. Black lines indicate nucleotide differences. C) Multiple sequence alignments between protein sequences of *tbAGT1, lgAGT1*, and *ncAGT1*. White gaps indicate amino acid differences.

Fig 2. Phylogeny of *Saccharomyces MALT* genes. ML phylogenetic tree of *MALT* genes described in *S. cerevisiae, S. eubayanus*, and lager-brewing hybrids. The scale bar equals the number of nucleotide substitutions per site. Black "*" indicate genes characterized as encoding proteins capable of transporting maltotriose. Gray "*" indicates genes encoding transporters whose ability to transport maltotriose is ambiguous.

Fig 3. Evolution and validation of the chimeric maltotriose transporter Malt434. A) After continuous culturing on maltotriose with a small amount of added glucose, yHKS210, which was originally unable to use maltotriose (MalTri-), evolved the ability to consume maltotriose (MalTri+). B) Strain yHEB1593, which is a backcross between yHKS210 and yHEB1505, was also MalTri+. C) To test the inheritance of maltotriose utilization, yHEB1593 was sporulated. The panel shows a subset of tetrads screened growing on SC + 2% maltotriose. Examples of MalTri- spores in Tetrad 1 are circled in red, and MalTri+ examples are circled in green. Whole genome sequencing of MalTri+ and MalTri- pools showed that maltotriose utilization perfectly correlated with the presence/absence of *MALT434*. D) Reciprocal hemizygosity test (78) of the *MALT4/MALT434* locus in the backcross strain yHEB1593. E) Table of initial and day-three

OD_{600} (OD) readings of yHKS210, yHEB1505, yHEB1593, yHEB1853, and yHEB1854 on SC + 2% maltotriose as the sole carbon source. N = 3, standard deviation in parentheses. * Control grown in SC + 0.04% glucose to reflect the approximate amount of growth expected from contamination with other carbon sources when using 98% pure maltotriose.

Fig 4. Sequence architecture of *MALT434*. A) Schematic of the origin of *MALT434*. B) Line graphs representing the identity between nucleotide sequences of *MALT3* and *MALT4* from yHKS210 to *MALT434* over 10-bp sliding windows. C-D) Segment of the alignment of the chimeric region between Malt3, Malt4, Malt434, scAgt1, and lgAgt1. The region highlighted in yellow in the Malt434 sequence indicates the chimeric region. The regions underlined with a red dashed line are predicted transmembrane domains. The amino acids highlighted in red are predicted maltose-binding residues. The residues highlighted in blue were experimentally found to be important for maltotriose transport by Smit *et al.* 2008.

Fig 5. Heterologous expression of *MALT434*. A) Evolution of non-maltotriose utilizing strain (MalTri-), yHKS210, to maltotriose utilizing (MalTri+) strain, yHEB1505, by serial passing on maltotriose containing media (same as Fig 3A). B) Insertion of *MALT434* into vector pBM5155 for doxycycline-inducible heterologous expression in MalTri- strains. C) Transformation of *MALT434* expression plasmid in MalTri- *S. eubayanus* strains yHKS210 and yHRVM108. D) Table of initial and day-six OD_{600} (OD) measurements of parent strains and strains carrying the *MALT434* expression plasmid grown in SC media with maltotriose as the sole carbon and doxycycline to induce plasmid expression. N = 3, standard deviation in parentheses. * Control grown in SC + 0.04% glucose + doxycycline to reflect the approximate amount of growth expected from contamination with other carbon sources when using 98% pure maltotriose.

Supporting Information

S1 Fig. Protein structural alignment between Malt3, Malt4, Malt434, scAgt1, and lgAgt1. Protein structural alignment between Malt3, Malt4, Malt434, scAgt1, and lgAgt1. The purple blocks represent predicted alpha helices, and the orange lines represent predicted beta strands. Red ticks mark predicted maltose-binding sites. Blue ticks mark residues found to be important for maltotriose transport by Smit *et al.* 2008. A green tick marks the location of the single non-synonymous substitution between *lgAGT1* and *tbAGT1*. Arrows point to alpha helices in Malt434 whose predicted sizes are reduced compared to other transporters in the alignment.

S1 Table. Mutation prediction analyses. Results of mutation prediction analyses for E18V, the sole amino acid substitution in the lgAgt1 protein sequence, relative to tbAgt1.

S2 Table. Maltose growth assay. Growth on maltose of single-colony isolates. Isolated from adaptive evolution of yHRVM108 on 2% maltose + 0.1% glucose. N = 3. * Control grown in SC + 0.04% glucose to reflect the approximate amount of growth expected from contamination with other carbon sources when using 98% pure maltotriose.

S3 Table. Maltotriose growth assay. Growth on maltotriose of single-colony isolates from adaptive evolution experiments. Strains were evolved with either maltotriose or maltose as the primary carbon source (2%) with 0.1% added glucose. N = 3. * Control grown in SC + 0.04% glucose to reflect the approximate amount of growth expected from contamination with other carbon sources when using 98% pure maltotriose.

S4 Table. Strains and plasmids used in this work.

S5 Table: Oligonucleotides used in this work.

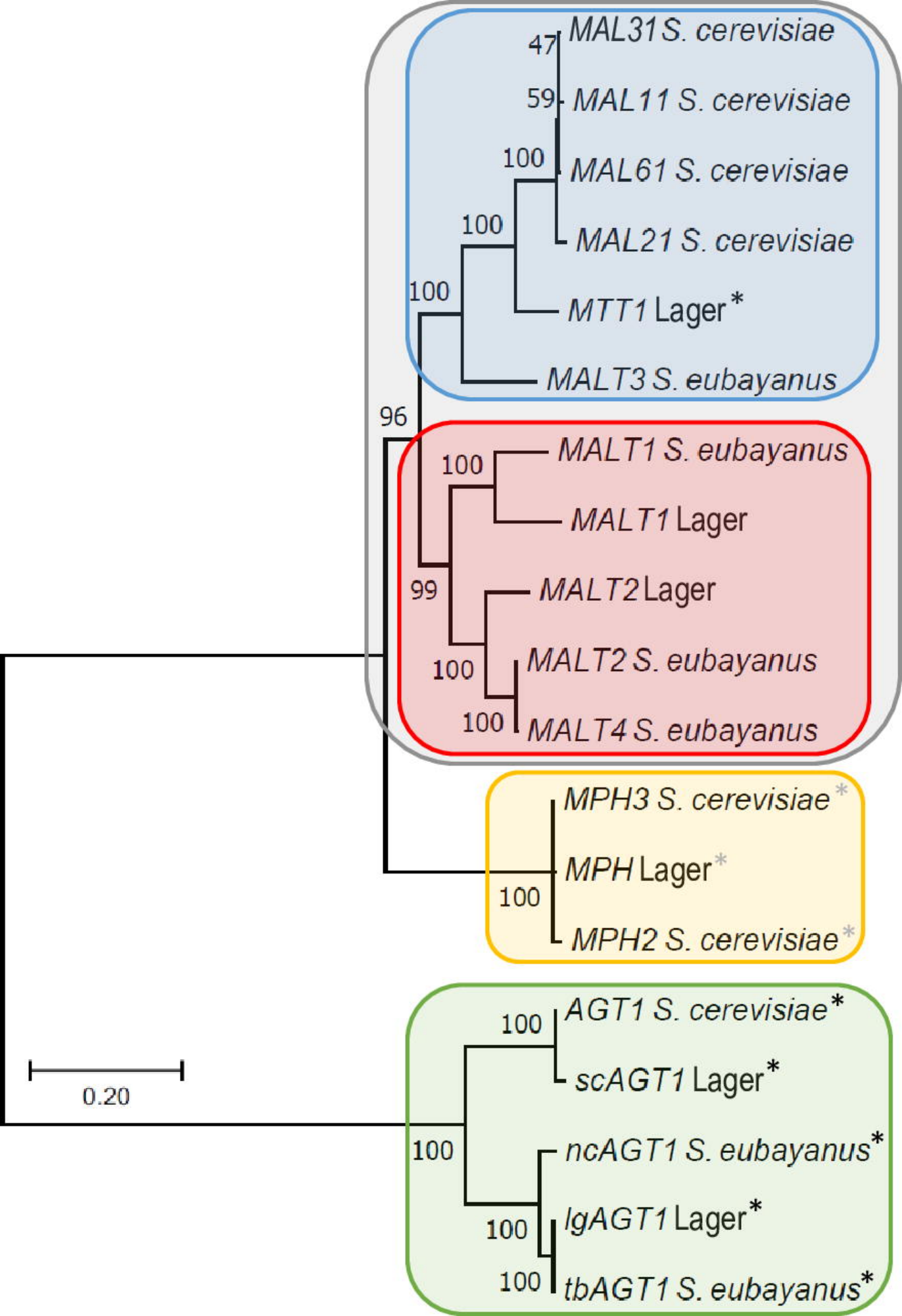
A.

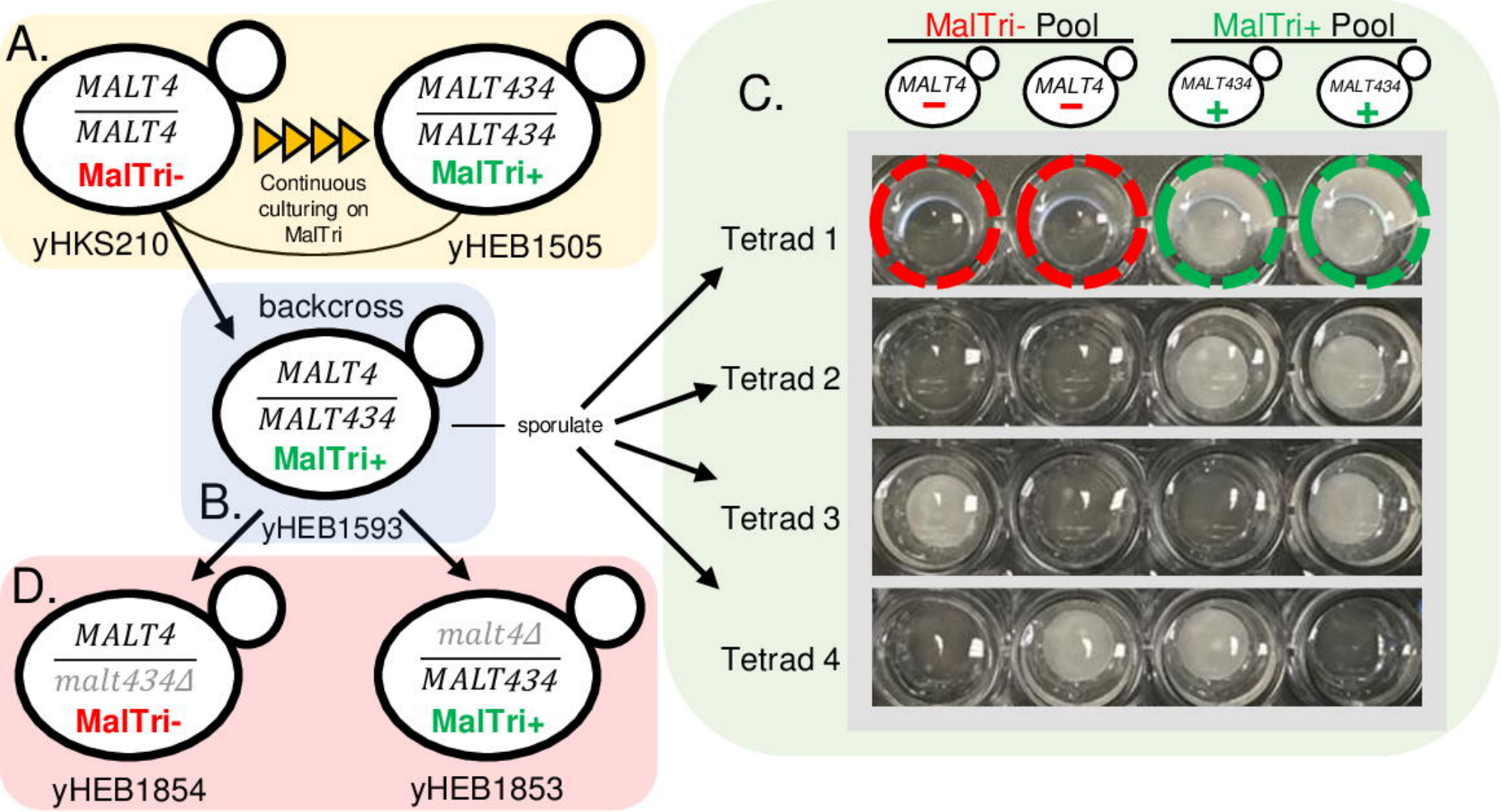
nuc	*lgAGT1*	*ncAGT1*	*scAGT1*	*AGT1*
tbAGT1	99.89	95.74	79.26	80.16
lgAGT1		95.63	79.15	80.05
ncAGT1			79.75	80.66
scAGT1				98.64

aa	lgAgt1	ncAgt1	scAgt1	Agt1
tbAgt1	99.84	95.57	83.74	86.07
lgAgt1		95.41	83.58	85.90
ncAgt1			83.25	85.41
scAgt1				97.24

B.

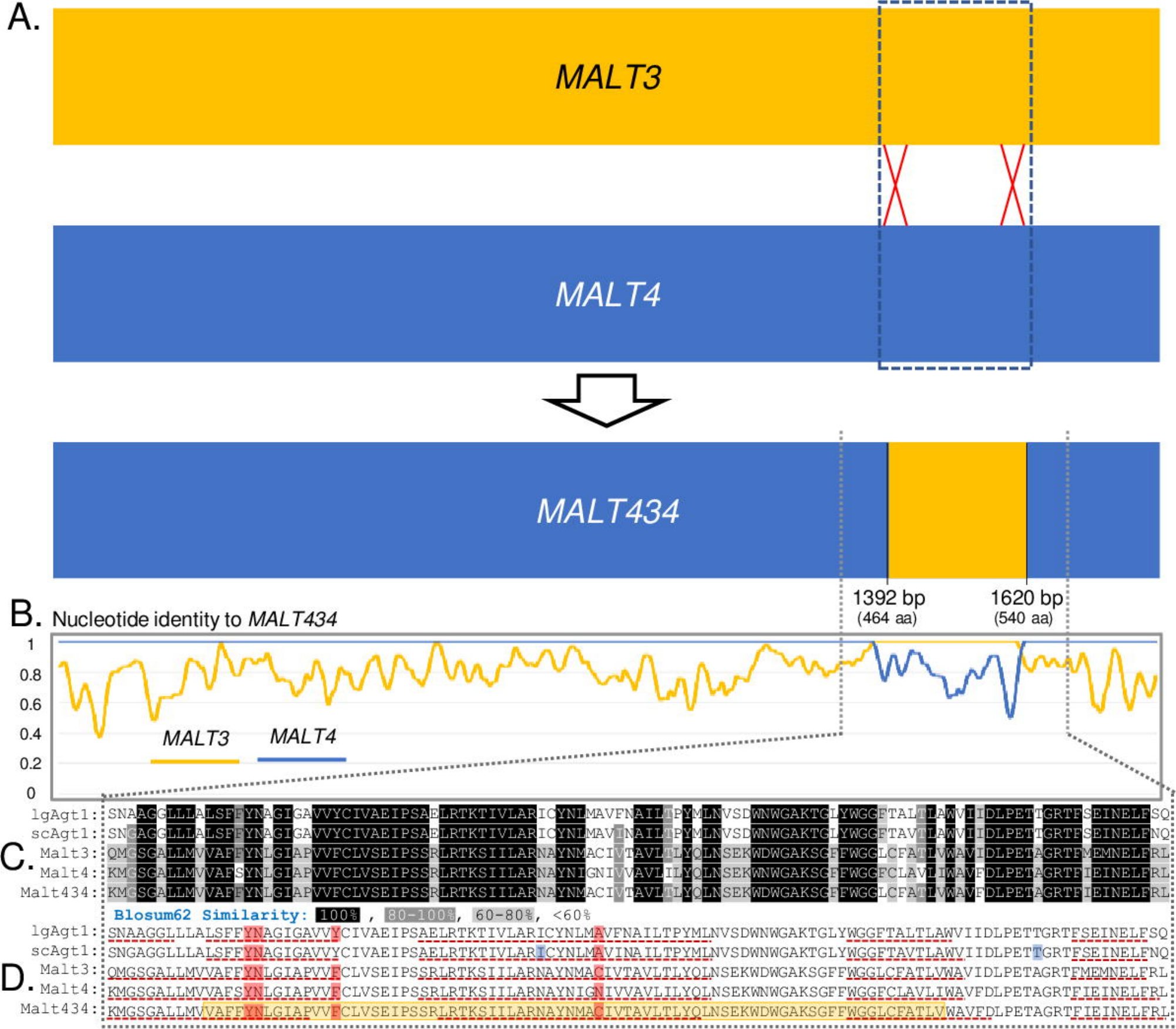
C.

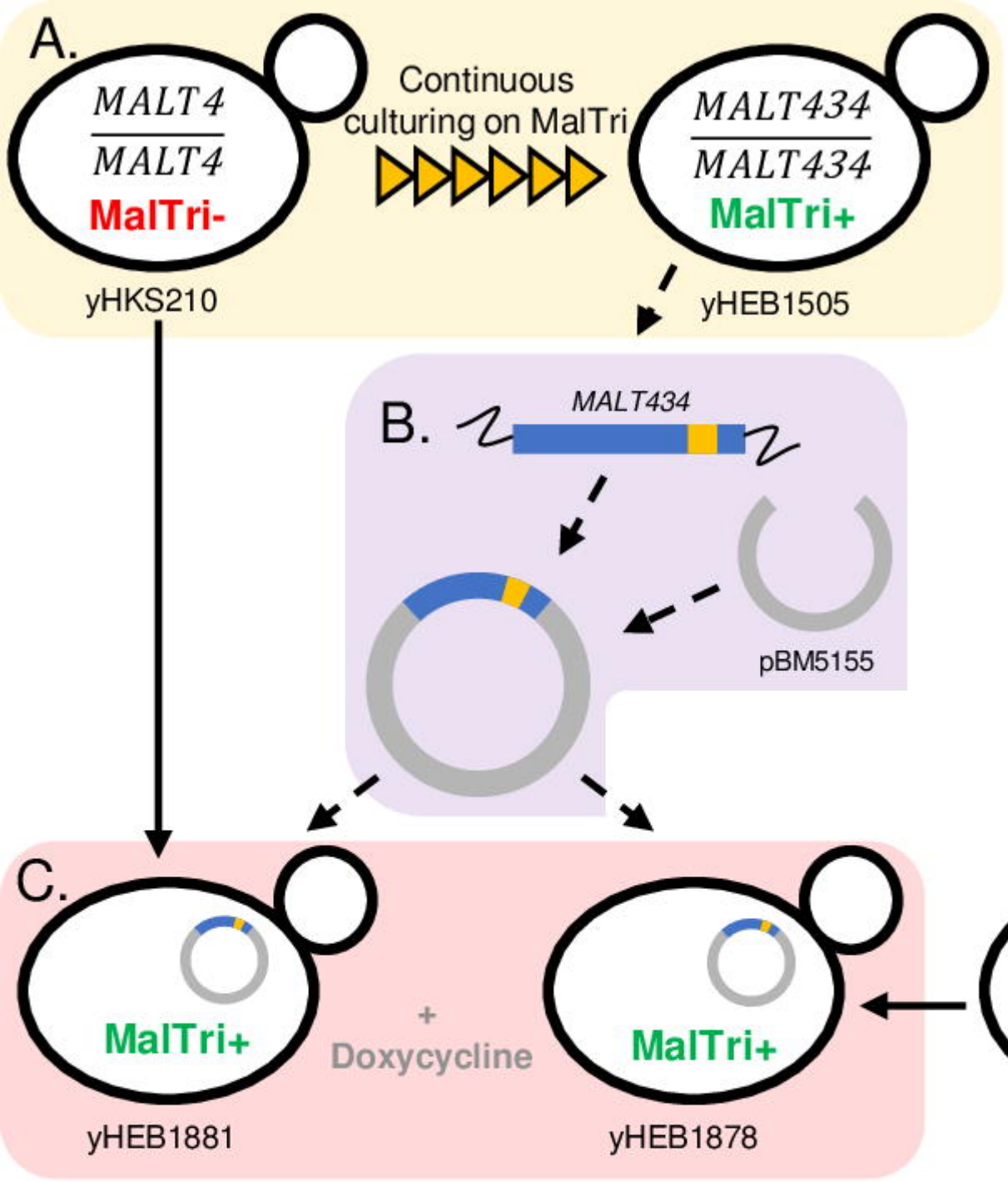

MAL31 S. cerevisiae
47
MAL11 S. cerevisiae
59
MAL61 S. cerevisiae
100
MAL21 S. cerevisiae
100
MTT1 Lager*
100
MALT3 S. eubayanus
96
MALT1 S. eubayanus
100
MALT1 Lager
MALT2 Lager
99
MALT2 S. eubayanus
100
MALT4 S. eubayanus
100
MPH3 S. cerevisiae*
MPH Lager*
100
MPH2 S. cerevisiae*
AGT1 S. cerevisiae*
100
scAGT1 Lager*
ncAGT1 S. eubayanus*
100
lgAGT1 Lager*
100
tbAGT1 S. eubayanus*
100
0.20



E.

Strain	Description	Initial OD	Day 3
yHKS210*	Wild admixture strain	0.22 (+/-0.01)	0.68 (+/-0.06)
yHKS210	Wild admixture strain	0.14 (+/-0.02)	0.39 (+/- 0.01)
yHEB1505	Single-colony isolate of yHKS210 evolved in maltotriose	0.16 (+/-0.02)	1.46 (+/- 0.01)
yHEB1593	yHKS210 x yHEB1505	0.12 (+/-0.01)	1.24 (+/- 0.02)
yHEB1854	yHEB1593 *MALT4/malt434Δ*	0.13 (+/-0.01)	0.39 (+/- 0.01)
yHEB1853	yHEB1593 *malt4Δ/MALT434*	0.11 (+/-0.01)	1.33 (+/- 0.04)


A.
MALT3
MALT4
MALT434
1392 bp
(464 aa)
1620 bp
(540 aa)
B.
Nucleotide identity to MALT434
1
0.8
0.6
0.4
0.2
0
MALT3
MALT4
C.
lgAgt1:
scAgt1:
Malt3:
Malt4:
Malt434:
Blosum62 Similarity: 100% , 80-100%, 60-80%, <60%
D.
lgAgt1:
scAgt1:
Malt3:
Malt4:
Malt434:



D.

Strain	Plasmid	Initial OD	Day 6
yHKS210*	-	0.19 (+/-0.04)	0.52 (+/-0.03)
yHKS210	-	0.14 (+/-0.02)	0.39 (+/-0.05)
yHEB1881	*MALT434*	0.11 (+/- 0.01)	1.31 (+/- 0.05)
yHRVM108*	-	0.16 (+/-0.05)	0.48 (+/-0.01)
yHRVM108	-	0.12 (+/-0.03)	0.46 (+/-0.03)
yHEB1878	*MALT434*	0.12 (+/-0.01)	1.46 (+/-0.04)