Specialization Restricts the Evolutionary Paths Available to Yeast Sugar Transporters

Johnathan G. Crandall (ID),[1] Xiaofan Zhou (ID),[2,3] Antonis Rokas (ID),[3] Chris Todd Hittinger (ID)[1,*]

[1]Laboratory of Genetics, J. F. Crow Institute for the Study of Evolution, Center for Genomic Science Innovation, DOE Great Lakes Bioenergy Research Center, Wisconsin Energy Institute, University of Wisconsin-Madison, Madison, WI 53726, USA

[2]Guangdong Province Key Laboratory of Microbial Signals and Disease Control, Integrative Microbiology Research Center, South China Agricultural University, Guangzhou 510642, China

[3]Department of Biological Sciences and Evolutionary Studies Initiative, Vanderbilt University, Nashville, TN 37235, USA

*Corresponding author: E-mail: cthittinger@wisc.edu.

Associate editor: Jeffrey Townsend

Abstract

Functional innovation at the protein level is a key source of evolutionary novelties. The constraints on functional innovations are likely to be highly specific in different proteins, which are shaped by their unique histories and the extent of global epistasis that arises from their structures and biochemistries. These contextual nuances in the sequence–function relationship have implications both for a basic understanding of the evolutionary process and for engineering proteins with desirable properties. Here, we have investigated the molecular basis of novel function in a model member of an ancient, conserved, and biotechnologically relevant protein family. These Major Facilitator Superfamily sugar porters are a functionally diverse group of proteins that are thought to be highly plastic and evolvable. By dissecting a recent evolutionary innovation in an α-glucoside transporter from the yeast *Saccharomyces eubayanus*, we show that the ability to transport a novel substrate requires high-order interactions between many protein regions and numerous specific residues proximal to the transport channel. To reconcile the functional diversity of this family with the constrained evolution of this model protein, we generated new, state-of-the-art genome annotations for 332 Saccharomycotina yeast species spanning ∼400 My of evolution. By integrating phylogenetic and phenotypic analyses across these species, we show that the model yeast α-glucoside transporters likely evolved from a multifunctional ancestor and became subfunctionalized. The accumulation of additive and epistatic substitutions likely entrenched this subfunction, which made the simultaneous acquisition of multiple interacting substitutions the only reasonably accessible path to novelty.

Key words: maltose, maltotriose, lager yeast, gene conversion, gene duplication, subfunctionalization, MFS, sugar transporter.

Introduction

Many key evolutionary innovations arise from changes to protein sequences that alter their function (Cheng 1998; Zhang et al. 2002; Clark et al. 2003; Dorus et al. 2004; Lunzer et al. 2005; Nielsen et al. 2005; Hoekstra et al. 2006; Christin et al. 2007; Yokoyama et al. 2008; Voordeckers et al. 2012; Projecto-Garcia et al. 2013; Kaltenbach et al. 2018; Jabłońska and Tawfik 2022). Occasionally, these changes stem from dramatic mutational events, including the creation of highly novel coding sequences by gene conversion or ectopic recombination resulting in chimeric proteins (Long and Langley 1993; Nurminsky et al. 1998; Wang et al. 2000; Long et al. 2003; Patthy 2003; Zhang et al. 2004; Ciccarelli et al. 2005; Arguello et al. 2006; Rogers et al. 2010; Rogers and Hartl 2012; Leffler et al. 2017; Méheust et al. 2018; Baker and

Hittinger 2019; Smithers et al. 2019; Brouwers et al. 2019b; Baker et al. 2022). While gene conversion can theoretically accelerate the rate of evolution (or even enable adaptation altogether) by bypassing deleterious intermediates, this effect is primarily attributable to the presence of a rugged fitness landscape (Kauffman and Levin 1987; HANSEN et al. 2000; Cui et al. 2002; Bittihn and Tsimring 2017). Such rugged landscapes are manifestations of epistasis in the genotypic combinations underlying the phenotypic map and are prevalent in some empirical systems (Wright 1931, 1932; Smith 1970; Townsend et al. 2003; Weinreich et al. 2005, 2006, 2013; Gong et al. 2013; De Visser and Krug 2014; Sarkisyan et al. 2016; Starr and Thornton 2016; Wu et al. 2016; Pokusaeva et al. 2019; Yi and Dean 2019; Nishikawa et al. 2021; Park et al. 2022; Meger et al. 2024; Metzger et al. 2024). For other proteins,

Open Access

the fitness landscape may be much smoother, meaning that stepwise mutations with additive effects can underlie functional evolution (Lunzer et al. 2005; Bridgham et al. 2006; Weinreich et al. 2006; Poelwijk et al. 2007; Campbell et al. 2016; Kaltenbach et al. 2018; Srikant et al. 2020). In cases where novel protein function is linked to gene conversion events between homologs, these observations therefore raise a fundamental question: are such dramatic mutational events required to evolve new function, or are they probabilistic shortcuts in the evolutionary process whose prevalence is a predictable function of their combined effect size and relative mutation rate? Answering this question has significant implications for understanding and predicting the most likely evolutionary trajectories, as well as for designing and engineering novel proteins with desirable functions.

Recently, several remarkably parallel cases of functional innovation have been linked directly or speculatively to gene conversion events in an ecologically and biotechnologically relevant protein family: maltose transporters in *Saccharomyces* yeasts (Baker and Hittinger 2019; Brouwers et al. 2019b; Hatanaka et al. 2022). This protein family consists of transporters similar to the *Saccharomyces cerevisiae* Mal31 protein, which has high specificity and high affinity for the disaccharide maltose, which contains two glucose moieties (Cheng and Michels 1991; Stambuk and Araujo 2001; Salema-Oom et al. 2005; Alves et al. 2008; Brown et al. 2010). Mal31-like proteins are encoded in nearly all genomes of *Saccharomyces* and some closely related species, and they are frequently encoded by multiple paralogs within each genome.

Maltose uptake is also mediated by a second family of proteins, which are related to *S. cerevisiae* Agt1. In contrast to the Mal31-like proteins, Agt1 is a generalist α-glucoside transporter with a broad substrate range, but it has generally lower affinity for those substrates (Han et al. 1995; Stambuk et al. 1999, 2000; Alves et al. 2008; Trichez et al. 2019). Notably, Agt1 can transport the glucose trisaccharide maltotriose, a molecule that is biochemically similar to maltose but contains a third glucose moiety. Although sometimes referred to as Mal11, Agt1 is a functionally distinct protein with ≈57% amino acid sequence identity to the Mal31-like proteins. In contrast to the Mal31-like proteins, Agt1-like proteins are rarer, both in presence and in paralog number, in the genomes of *Saccharomyces* yeasts and close relatives (Duval et al. 2010; Horák 2013).

The α-glucoside transporters (Agts) of *Saccharomyces* include the Agt1-like ("generalist") and Mal31-like (high-specificity) proteins, as well as Mph2/3-like proteins (Day et al. 2002), which also have high specificity, albeit for the α-glucoside turanose (Brown et al. 2010). These Agts have been extensively studied due to their important role in the production of beer. Maltose and maltotriose are the two most abundant sugars in brewer's wort (Meussdorfer and Zarnkow 2009), and their transport into the cell is the rate-limiting step in their fermentation (Zastrow et al. 2001; Horák 2013). The rarity of maltotriose transporters, such as Agt1, which almost always manifests

as an inability to ferment this carbon source, therefore presents a barrier to the use of many nondomesticated yeasts in brewing applications.

This barrier is exemplified in *Saccharomyces eubayanus*, the wild, cold-tolerant parent of industrial lager-brewing hybrids (Libkind et al. 2011), whose development for commercial brewing is of great interest (Gibson et al. 2017; Hittinger et al. 2018; Cubillos et al. 2019). As almost all strains of *S. eubayanus* lack generalist Agts capable of transporting maltotriose (Brickwedde et al. 2018; Brouwers et al. 2019a; Bergin et al. 2022), multiple attempts have been made to evolve maltotriose transporters de novo in *S. eubayanus* strains, using both mutagenesis (Brouwers et al. 2019b) and adaptive laboratory evolution (Baker and Hittinger 2019). These experiments, performed independently in different backgrounds of *S. eubayanus*, yielded results that were as remarkable in their similarity as they were unexpected. In both cases, ectopic gene conversion between paralogous high-specificity (Mal31-like) maltose transporters without any native maltotriose transport capacity (Brickwedde et al. 2018; Baker and Hittinger 2019) resulted in chimeric proteins capable of transporting maltotriose.

Lending weight to the notion that recombination may be a common mechanism by which transporters in the high-specificity Agt family evolve new function, two newly discovered *S. cerevisiae* transporters (Hatanaka et al. 2022), as well as the Mty1 protein (Dietvorst et al. 2005; Salema-Oom et al. 2005), may possess signatures of more ancient gene conversion events (Brouwers et al. 2019b). All these proteins transport maltotriose, but they cluster with Mal31-like proteins in phylogenetic analyses (Baker and Hittinger 2019; Hatanaka et al. 2022). Nonetheless, it remains unclear whether these dramatic mutational events are required for the evolution of novel function in this family or whether they are simply enriched due to the dynamic nature of the subtelomeric regions in which these genes reside (Mefford and Trask 2002; Fairhead and Dujon 2006; Gordon et al. 2009; Brown et al. 2010; Yue et al. 2017; Peter et al. 2018; Liu et al. 2019; O'Donnell et al. 2023).

The yeast α-glucoside transporters are H$^+$ symporters belonging to the sugar porter family (TCDB: 2.A.1.1) of the Major Facilitator Superfamily (MFS), a vast, ubiquitous, and ancient group of transmembrane proteins present in all domains of life (Marger and Saier 1993; Pao et al. 1998; Saier 2000; Wang et al. 2020; Saier et al. 2021). Across great evolutionary distances, sugar porters share the highly characteristic MFS fold consisting of 12 transmembrane helices (TMHs) surrounding a hydrophilic central cavity that constitutes the transport channel (Abramson et al. 2003; Guan and Kaback 2006; Sun et al. 2012; Deng et al. 2014; Quistgaard et al. 2016; Bosshart and Fotiadis 2019; Kaback and Guan 2019; Paulsen et al. 2019; Drew et al. 2021). These TMHs are organized into two pseudosymmetrical six-helix bundles (N- and C-terminal), which are separated by a long intracellular helical linker (ICH domain). The transport channel is

surrounded by four helices from each bundle, and TMHs stack tightly against their intrabundle partners, with additional contacts between the N- and C-terminal domains at the interbundle interface. In *S. cerevisiae* Agt1, the sugar substrate and/or proton are thought to be bound primarily by charged residues projecting into this central cavity, which are conserved across fungal Agts (Henderson and Poolman 2017; Trichez et al. 2019). More generally, substrate affinity and specificity in MFS sugar transporters are mediated by extensive hydrogen bonding and occasionally by hydrophobic interactions between the sugar and the protein, as well as steric constraints that limit substrate accommodation; moreover, there is a growing appreciation for the fine-scale and occasionally cryptic contributions to affinity by residues within Van der Waals distance of the substrate (Kasahara et al. 1997; Kasahara and Kasahara 1998, 2000, 2010; Guan and Kaback 2006; Kasahara et al. 2006, 2007, 2009, 2011; Guan et al. 2007; Sun et al. 2012; Deng et al. 2014; Farwick et al. 2014; Deng et al. 2015; Bosshart and Fotiadis 2019; Kaback and Guan 2019; Drew et al. 2021; Guan and Hariharan 2021).

Nonetheless, the extensive and exquisite biochemical study of MFS sugar transporters has almost exclusively focused on the determinants of native substrate binding and affinity in extant proteins, while questions about how such proteins could evolve the capacity to transport a novel substrate de novo have been largely unaddressed. Understanding evolution-informed design principles in this protein family could enable the engineering of desirable properties in tractable proteins, with significant implications for industrial processes, including the fermentation of cellulosic and hemicellulosic biomass into next-generation biofuels and bioproducts (Ha et al. 2013; Farwick et al. 2014; Young et al. 2014; Turner et al. 2016; Hara et al. 2017; Oh et al. 2017; Casa-Villegas et al. 2018; Kim et al. 2018; Nijland et al. 2018; de Ruijter et al. 2020; Nijland and Driessen 2020; Oh and Jin 2020).

To this end, we aimed to dissect the molecular genetic basis of novel function in the chimeric *S. eubayanus* maltotriose transporter MalT434. *MALT434* arose from an ectopic gene conversion event between genes encoding two paralogous maltose transporters, MalT3 and MalT4, which resulted in the replacement of ~230 bp of the *MALT4* gene with the homologous portion of *MALT3* (Baker and Hittinger 2019). Both MalT3 and MalT4 are members of the high-specificity maltose transporter family and incapable of transporting maltotriose (Brickwedde et al. 2018; Baker and Hittinger 2019), suggesting that intramolecular epistasis between their protein regions underlies the emergent maltotriose transport by MalT434. The translocated region of *MALT3* encodes TMH 11 and portions of TMHs 10 and 12 (Fig. 1a), and it introduced 11 nonsynonymous mutations to the protein-coding sequence of *MALT4* (Fig. 1b). All three proteins are predicted to have virtually identical structures across their entire folds [pairwise root mean square deviation (RMSD) = 0.955 Å] and TMHs 10 to 12 (0.909 Å, supplementary fig.

S1, Supplementary Material online), suggesting that novel substrate transport might stem from a specific combination of substrate-interacting residues from distal protein regions in MalT434, rather than a global change to protein structure. In the simplest model, as few as a single interacting residue from each protein region could underlie the emergence of novel function, which would make the evolution of new function in this family predictable and tunable; in the most complex model, all 120 amino acid differences between the 2 parental transporters could contribute, which would render the evolution of new function incredibly difficult.

Here, we show that the basis of maltotriose transport is remarkably complex in this model neofunctionalized transporter. Novel function is shaped by a combination of additive and nonadditive interactions between as many as seven regions in the MalT4 backbone and six substitutions across TMHs 10 and 11. At one critical site, very few amino acids can support novel function, which further limits the evolutionary paths available to the wild-type protein; at other sites, these requirements are less stringent. We propose that, overall, novel substrate transport is enabled by widening the transport channel while simultaneously creating a favorable electrostatic environment for the bulkier trisaccharide molecule. Finally, we reconstruct the evolutionary history of the high-specificity and generalist yeast Agts and their relationships to other sugar porters; unexpectedly, we show that the specialist maltose transporters are likely derived and subfunctionalized from a generalist ancestor. This specialization likely involved a gradual refinement of the transport channel to specifically accommodate maltose with higher affinity, which makes the reacquisition of ancestral generalist function difficult to achieve. While our results indicate that rational engineering for novel substrate transport in this protein family is likely to be difficult, they also highlight the abundance and diversity of transporters in biotechnologically relevant yeast species, which could be readily mined for desirable functions that have been exquisitely refined over billions of years of evolution, as well as perhaps recombined into new functions.

Results

High-order Intramolecular Interactions Are Required to Evolve a Novel Function in Maltose Transporters

We first investigated the scope and complexity of intramolecular interactions shaping the emergence of novel function in MalT434. We defined functional protein units as the 12 TMHs, the ICH domain, and the partially unstructured intracellular N- and C-terminal regions, which were chosen based on these units' separable roles in substrate binding and translocation. We iteratively constructed novel chimeric genes encoding transporters from these MalT3 and MalT4 components and tested their ability to support growth on maltotriose when expressed from the native *MALT4* locus (Fig. 2). Unsurprisingly, the
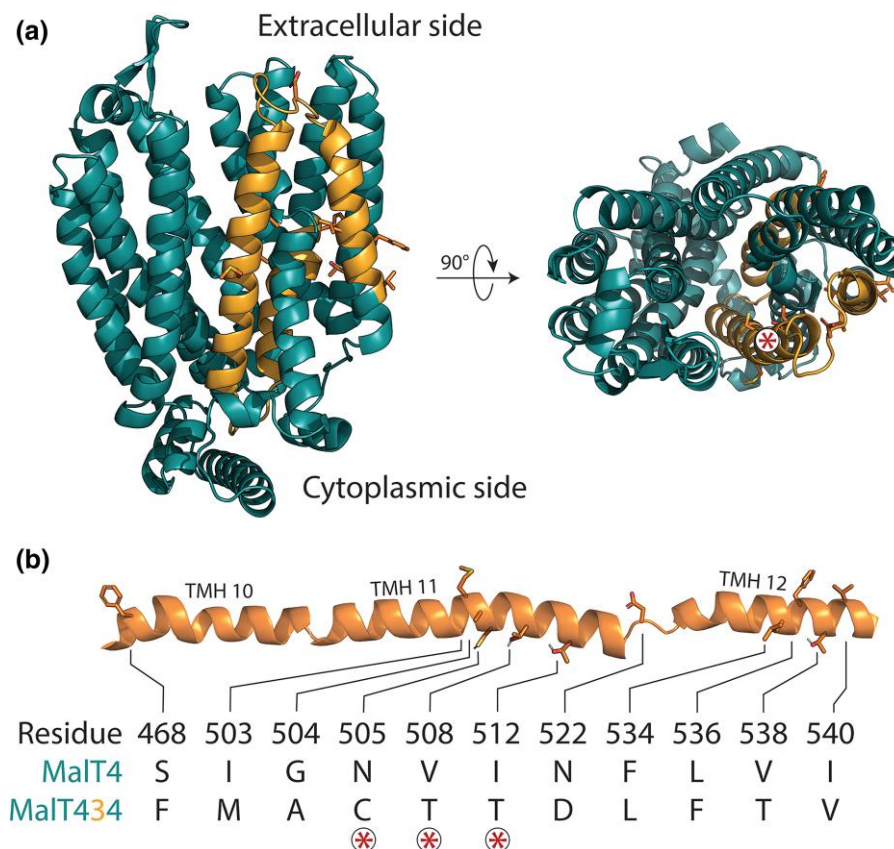
Fig. 1. Architecture of a chimeric neofunctionalized α-glucoside transporter. a) A structural model of the chimeric transporter MalT434 is shown from the side and top views, with alternating colors demarking regions contributed by different parental proteins. The top view is orientated looking down the transport channel. MalT3 side chains are drawn for the 11 substitutions between MalT4 and MalT434. The asterisk label marks the position of the three substitutions on a helical face that bounds the transport channel. b) Schematic of mutations. The 11 substitutions between MalT4 and MalT434 are drawn as side chains along the cartoon secondary structure of the protein, with loops that connect TMHs truncated for clarity. Polar hydrogens are shown. Asterisks mark the amino acids that face the transport channel.

C-terminal portion of MalT4 present in MalT434 was neither necessary (construct 1) nor sufficient (construct 17) for maltotriose transport; indeed, its replacement with the corresponding region of MalT3 improved growth on maltotriose by 15.3% ($P = 5.3 \times 10^{-4}$, Mann–Whitney U test). By contrast, replacement of TMHs 8 and 9 and the N-terminal half of TMH 10 with their MalT3 counterparts (construct 2) reduced growth by 11.6% compared with MalT434 ($P = 0.184$), while still supporting robust growth. Dissection of the region N-terminal to TMH 8 revealed that the key interaction enabling maltotriose transport occurs between one or more of TMHs 10 to 12 of MalT3 and TMH 7 from MalT4. While necessary, this region alone was not sufficient to enable maltotriose transport in every protein context. In addition to the epistatic interaction between TMHs 7 and 10 to 12, growth on maltotriose required the presence of TMHs 1 and 2 from MalT4 in combination with the ICH domain from MalT3 (construct 7), or alternatively, one or more of TMH 5, TMH 6, and the ICH domain from MalT4 (construct 15).

For chimeric constructs containing potentiating sequences at TMHs 5 to 7 and 10 to 12, growth on maltotriose generally increased additively with the number of MalT4 regions incorporated (linear regression, $P < 2.2 \times 10^{-16}$).

Nonetheless, we found significant support (analysis of variance, $P < 2.2 \times 10^{-16}$) for pairwise epistasis between the tested protein regions, including in the sign of the effects of the ICH domain and the C-terminal region (residues 541 to 613). For example, the addition of TMH 3 and TMH 4 from MalT4 in conjunction with MalT4 TMH 7 only increased growth on maltotriose if TMH 5 and TMH 6 from MalT4 were also present; similarly, the addition of TMH 1, TMH 2, and the ICH domain from MalT4 in conjunction with TMH 7 did not improve growth (construct 6 vs. 16, Fig. 2) unless in the presence of TMHs 3 to 6 from MalT4 (construct 2 vs. 13, 52% increase, $P = 2.4 \times 10^{-4}$). Along the quantitative functional spectrum of MalT3/4 chimeric proteins enabling growth on maltotriose, we therefore detected a complex combination of additive and epistatic intramolecular interactions among at least six protein regions.

Numerous Substitutions Are Required to Evolve a Novel Function in Maltose Transporters

We next dissected the contributions of the 11 substitutions in MalT434 relative to MalT4 (Fig. 1b) by introducing subsets of these to the gene encoding the native MalT4
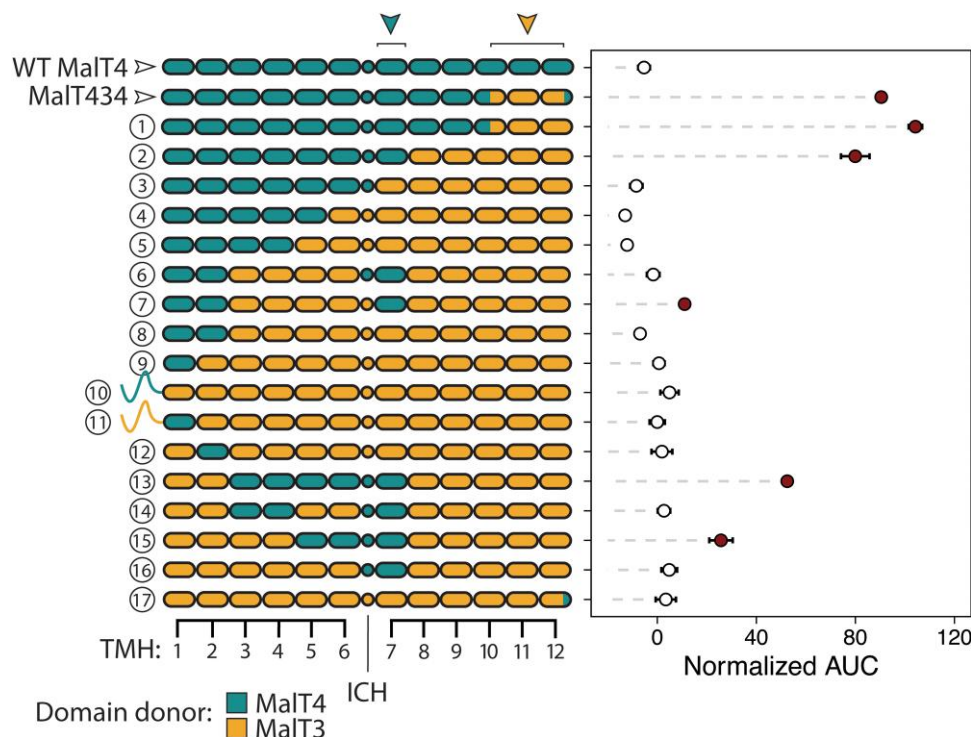
Fig. 2. High-order intramolecular interactions are required to evolve a novel function in chimeric α-glucoside transporters. Points and bars show mean ± standard error of the mean of normalized growth on maltotriose (AUC) of strains expressing chimeric transporters or wild-type MalT4 (top row). Filled circles denote growth significantly greater than the negative control (P < 0.01, Mann–Whitney U test with Benjamini–Hochberg correction). The architecture of each tested transporter is depicted as a cartoon on the y axis, where rounded rectangles represent each of the 12 TMHs and circles represent the ICH domain that links the N- and C-terminal six-helix bundles; regions are colored by parental protein identity. In almost every case, the N- and C-terminal intracellular regions have the same parental protein identity as the neighboring TMH and are omitted for clarity; the two exceptions are depicted. Inverted arrows and brackets indicate the location and identity of protein regions underlying the largest detected intramolecular interaction.

protein (Fig. 3). We first tested the effect of a pair of suggestive substitutions, S468F and N522D, which were both unique in their location in the tertiary structure and differed notably in side chain chemistry (protein 14, Fig. 3). Nonetheless, this pair of substitutions was insufficient for novel function in MalT4, so we coarsely tested the effect of the sets of substitutions occurring before and after the end of TMH 11. Introduction of the five substitutions between residues 522 and 540, which span an extracellular loop and the majority of TMH 12, was insufficient to confer any growth on maltotriose (protein 2, Fig. 3). By contrast, the six mutations affecting TMHs 10 and 11 were sufficient to confer growth on maltotriose (protein 1) and even improved it by 13.3% relative to MalT434 (P = 5.6 × 10^{-7}, Mann–Whitney U test). Within this contiguous patch of substitutions, however, the contribution of individual amino acids to novel function was remarkably complex. Reversion of the six mutations singly to their MalT4 identity revealed that each had a significant effect on maltotriose growth, ranging from a 23.5% reduction (A504G, protein 5; P = 2 × 10^{-6}) to its complete abrogation (C505N, protein 6; P = 5.2 × 10^{-11}), with an average effect of 57.1%. We detected significant (P < 2.2 × 10^{-16}) evidence of pairwise epistasis between substitutions, regardless of whether we considered all 11 sites or only the 6

on TMHs 10 and 11. Epistatic effects were notably non-uniform among tested combinations: for example, two single reversion mutations (M503I/protein 4 and T508V/protein 7) had similar effects of 49.1% (P = 3.2 × 10^{-7}) and 44.1% (P = 9.1 × 10^{-13}) when introduced in the 6-substitution background that supported robust growth on maltotriose. By contrast, when introduced in a four-substitution background with reduced ability to support growth on maltotriose (protein 9), the effect of M503I remained large (42.6%, protein 11; P = 0.002), while T508V effected only a small further reduction (4.97%, protein 10; P = 0.8). Overall, we found that establishing novel function in MalT4 required a combination of three amino acid substitutions only accessible through a minimum of four nonconsecutive nucleotide substitutions to the wild-type gene: N505C (two nucleotide substitutions), I512T (one substitution), and one of I503M (one substitution) or V508T (two substitutions).

Granular Mapping of Epistasis Between Distal Protein Regions

Given the size of interacting protein regions and the complexity of their contributions to novel function, we sought to identify the key difference in amino acid sequence
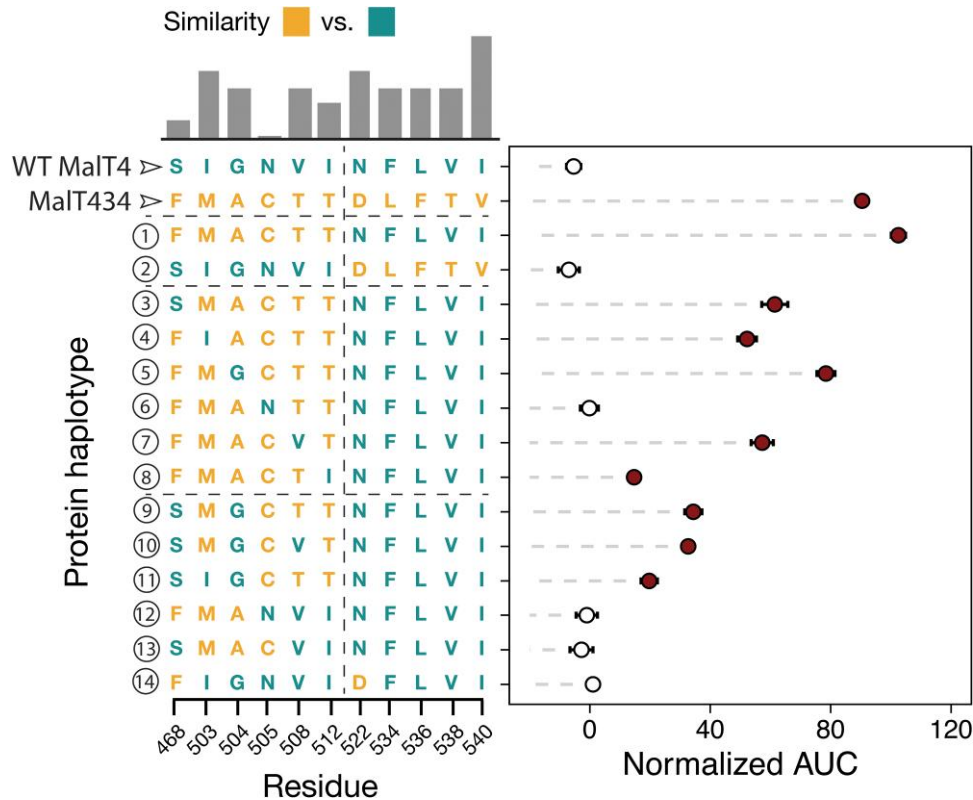
Fig. 3. Numerous substitutions are required to evolve a novel function in a maltose transporter. Points and bars show mean ± standard error of the mean of normalized growth on maltotriose (AUC) of strains expressing MalT4 variants. The genotype of each protein at the 11 sites that differ between MalT4 (top row) and MalT434 (second from top row) is depicted on the y axis. Filled circles denote growth significantly greater than the negative control ($P < 0.01$, Mann–Whitney U test with Benjamini–Hochberg correction). The bar chart shows rescaled BLOSUM similarity between the MalT4 and MalT3 residues at that site, with a higher bar indicating a more conservative substitution. Horizontal dotted lines in the protein haplotype grid separate related groups of genotypes. The vertical dotted line demarcates the substitutions that are sufficient (left) to impart novel function to MalT4 and those that are insufficient (right).

responsible for the large epistatic effect of TMH 7. The two parental transporters differ at six sites along TMH 7 (supplementary fig. S2a, Supplementary Material online): two neighboring substitutions (K357C and V358I, expressed relative to MalT4) occur at the intracellular C-terminal end, while two (A371I and V375T) are located approximately halfway along the helix and likely to be embedded in the plasma membrane. Two (A378T and S379Q) project into or neighbor the transport channel, differ in size and/or polarity, and are in close 3D proximity to mutated residues on TMH 11 in MalT434 (Fig. 4a; supplementary fig. S2b, Supplementary Material online). We reasoned that one or both of A378T and S379Q might have a large effect on the interaction between TMH 7 and the translocated region of MalT3 present in functional chimeric transporters. To test these hypotheses, we mutated each of these residues to their MalT3 identity, singly and in combination, in a gene encoding the MalT4 transporter harboring the six mutations on TMHs 10 and 11 that conferred maximal maltotriose transport (Fig. 4b). While the A378T mutation did not affect growth on maltotriose, S379Q abolished it completely. The large epistatic interaction between TMH 7 and TMH 11 can thus be attributed to a single amino acid.

Novel Transporter Function Is Constrained by Specific Biochemical Requirements and Context

The mutational event that generated MalT434, as well as our experiments dissecting it, only sampled variation between two binary states: the specific amino acid residues of the parental proteins at each homologous site. In native contexts, however, many more amino acid substitutions are accessible in mutational space through single- or multi-nucleotide mutations; for example, seven amino acid substitutions require only a single nucleotide change from an asparagine codon, which is the wild-type amino acid at the crucial 505 site. While we found complex interactions between many sites to contribute to novel function in MalT4, the evolution of maltotriose transport would be far less constrained and more accessible through sequential point mutations if biochemically similar amino acids at key sites could enable a degree of novel function because it would increase the mutational target size and pool of mutations conferring a fitness benefit (Miyazaki and Arnold 1999; Podgornaia and Laub 2015).

We thus sought to clarify the biochemical requirements for maltotriose transport in a specific potentiated context: a MalT4 transporter harboring S379, F468, M503, A504, T508, and T512. In this state, amino acid identity at position
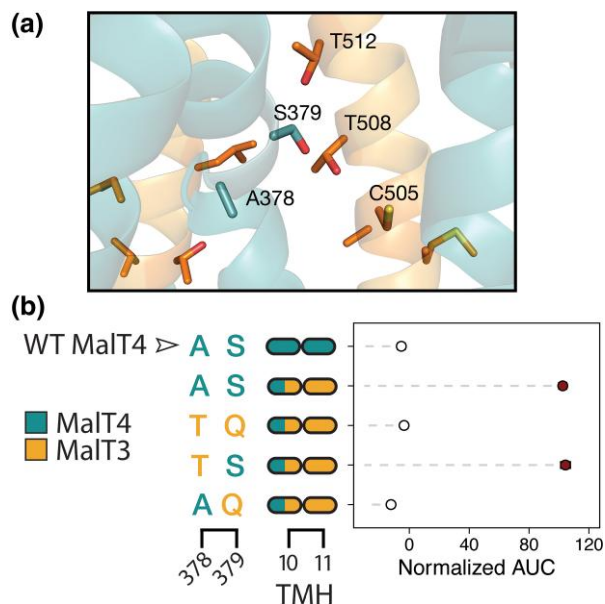
Fig. 4. A single amino acid underlies a large epistatic effect. a) Structural model of MalT434 with helices colored according to parental protein. Side chains are drawn for amino acids on TMHs 7, 11, and 12 that are polymorphic between MalT3 and MalT4, and those that are proximal to or project into the transport channel are labeled. b) Points and bars show mean ± standard error of the mean of normalized growth on maltotriose (AUC) of strains expressing transporter variants. Filled circles denote growth significantly greater than the negative control ($P < 0.01$, Mann–Whitney U test with Benjamini–Hochberg correction). Colors indicate parental protein identity at TMHs 10 and 11 (filled rectangular ovals) and residues 378 and 379 in TMH 7.

505 is crucial with the wild-type asparagine incapable of supporting growth on maltotriose and the recombinant cysteine supporting robust growth (Fig. 3). We successfully mutated this residue to 17 of the 20 possible amino acids, measured their ability to support growth on maltotriose, and used regression analyses to estimate the effect of side chain physicochemical properties on measured function (Fig. 5). Remarkably, only three substitutions supported any degree of statistically significant growth above baseline: serine, glycine, and cysteine. Side chain aromaticity, volume, composition, and hydropathy were all significant ($P \ll 0.01$) predictors of function, as was overall similarity to the wild-type residue asparagine. Even so, the strengths of these associations were almost entirely driven by the C505 variant: when these data were omitted, the global explanatory power was reduced dramatically (adjusted R^2: 0.2263 vs. 0.8664; F-statistic: 9.533 vs. 242). Although some physicochemical properties remained statistically significant predictors of function, the strengths of these associations were generally weak (maximum|Kendall's T|: 0.212).

Qualitatively, the fine-scale stringency of physicochemical requirements at position 505 was also noteworthy. Glycine, serine, and cysteine are three of the smallest amino acids, but amino acids with similar side chain volumes did not support growth on maltotriose. Serine and cysteine have side chains of similar size and structure capable of forming

hydrogen bonds, but they differ in their polarity and hydrophobicity; nonetheless, residues similar to cysteine in both of these metrics did not support novel function. Indeed, C505's ability to support novel function appeared to be the result of the specific combination of cysteine's physicochemical properties (supplementary fig. S3, Supplementary Material online), albeit not due to its unique capacity to form disulfide bridges (Drew et al. 2021). Remarkably, this effect was dependent on positional context within the transporter: while substituting cysteine to serine at 505 reduced growth by 71.2% ($P = 8.8 \times 10^{-5}$), making the orthogonal serine to cysteine substitution at another key site, S379 (Fig. 4) reduced growth by 17.7% ($P = 1.9 \times 10^{-6}$) while still supporting robust growth (supplementary fig. S4, Supplementary Material online). Thus, while serine was largely unable to recapitulate the effect of cysteine at 505, the similarity between the two was sufficient to satisfy the requirements for novel function at position 379. The same was not true of two other hydrogen bond–competent residues, glutamic acid and glutamine, whose introduction at position 379 abolished growth (supplementary fig. S4, Supplementary Material online). This result suggests that, while serine and cysteine are interchangeable at this site, interactions between physical and chemical side chain properties still play a role. Finally, we found further evidence for these fine-scale requirements at position 512, where mutation of the permissive threonine to valine reduced growth by 34.5% ($P = 7.4 \times 10^{-9}$), while still supporting significantly improved growth over the wild-type MalT4 residue isoleucine (78.1% increase, $P = 1.2 \times 10^{-6}$). In summary, we find that the strengths, stringencies, and bases of physicochemical requirements all vary between sites that are critical for establishing novel function in MalT434. These results suggest that the serendipitous acquisition of a set of epistatically sufficient residues is highly improbable by point mutations alone (Lynch 2005).

High-Specificity Transporters Are Evolutionarily Derived

The sum of our molecular analyses suggested that the acquisition of novel substrate transport by the high-specificity maltose transporter MalT4 is highly improbable and accessible only through the simultaneous acquisition of numerous interacting substitutions. This observation is consistent with previous failed attempts to establish a maltotriose transporter by introducing as many as 14 rational mutations to *S. cerevisiae* Mal61 (Hatanaka et al. 2022), a prototypical high-specificity maltose transporter closely related to MalT4. However, the presence of closely related generalist α-glucoside transporters, as typified by *S. cerevisiae* Agt1, suggests that this ability evolved at least once among yeast α-glucoside transporters. We sought to clarify the timing and mode of this historical evolutionary innovation by examining the phylogenetic relationships between the generalist and specialist α-glucoside transporters within Saccharomycotina yeasts, which have previously been assessed on only a few taxa (Brown et al. 2010; Cousseau et al. 2013; Baker and Hittinger 2019; de Ruijter et al. 2020; Hatanaka et al. 2022; Donzella et al. 2023).
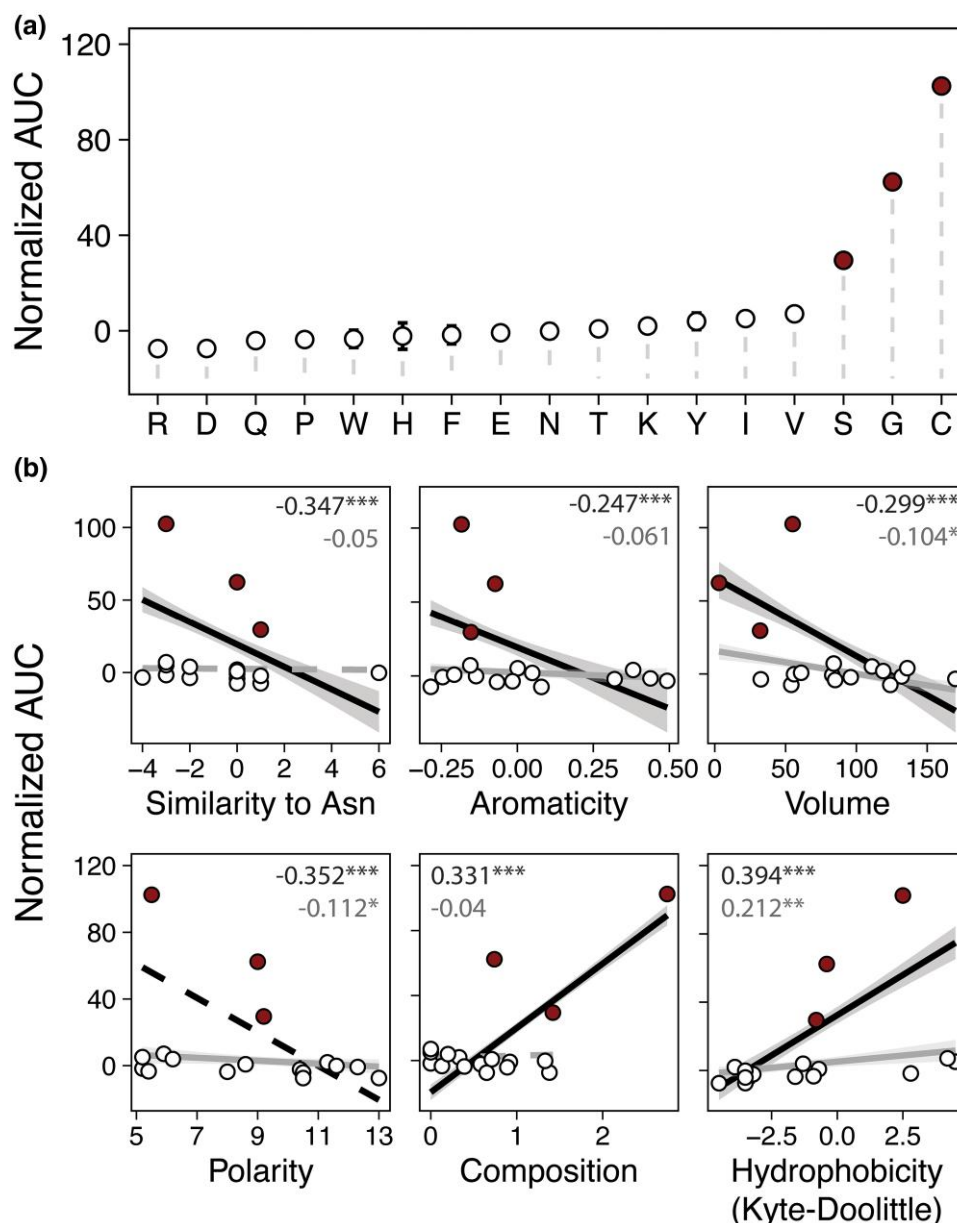
Fig. 5. Physicochemical requirements constrain the evolution of novel function. a) Points and bars show mean ± standard error of the mean of normalized growth on maltotriose (AUC) of strains expressing MalT4 variants. The x axis shows the amino acid identity at position 505; all variants share F468, M503, A504, T508, and T512. Filled circles denote growth significantly greater than the negative control ($P < 0.01$, Mann–Whitney U test with Benjamini–Hochberg correction). b) Correlations between growth and properties of the amino acid variant at position 505. Growth is plotted as in a) against physicochemical property or overall similarity to the wild-type residue at position 505, asparagine. Lines and shaded ranges show regressions and 95% confidence intervals for significant ($P < 0.05$) regressions for all data (black) or after removing observations for C505 (gray). Dotted lines show regressions that are not statistically significant. Inset text shows Kendall's τ; ***$P < 10^{-6}$, **$P < 10^{-4}$, and *$P < 0.05$.

We first generated high-quality protein-coding gene annotations for published genomes from 332 yeast species from the model subphylum Saccharomycotina, which spans more than 400 My of evolution (Shen et al. 2018). To formally test the expected monophyly of the α-glucoside transporters within the broader sugar porter family, we retrieved homologs of *S. cerevisiae* sugar porters from these predicted proteomes and constructed a comprehensive phylogeny of these 8,403 ecologically and biotechnologically relevant MFS proteins. This phylogeny split into several major clades, many of which contained

at least one functionally characterized protein from *S. cerevisiae* or another species (supplementary fig. S5, Supplementary Material online). Both the high-specificity (Mal31- and Mph2/3-like) and generalist (Agt1-like) α-glucoside transporters clustered in a monophyletic group (Agt clade) that excluded other sugar porter families. All proteins in the Agt clade from the newly circumscribed order Saccharomycetales (Groenewald et al. 2023) grouped together with strong support (Fig. 6a). The monophyly of the Saccharomycetales Agts was interrupted in two cases: (i) a single protein from *Ogataea naganishii* sister
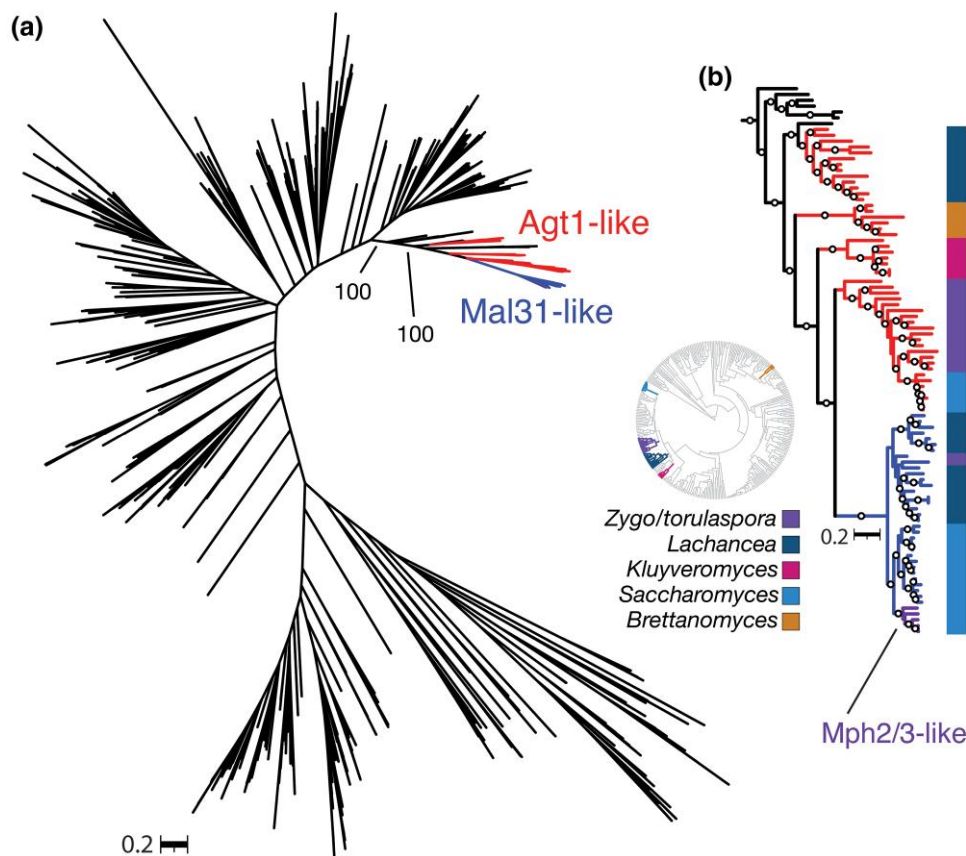
Fig. 6. The high-specificity maltose transporters are evolutionarily derived and restricted to a subset of Saccharomycetales. a) Unrooted consensus phylogeny of the α-glucoside transporter clade from 332 budding yeast genomes. Agt1-like and Mal31-like proteins from all Saccharomycetales are colored, as is the *Saccharomyces*-specific Mph2/3 clade. Bootstrap support is shown for two splits leading to the Saccharomycetales. b) Rooted consensus tree of the clade containing Saccharomycetales α-glucoside transporters. Branches are colored as in a) with the inclusion of a well-supported clade of *Brettanomyces* Agt1-like proteins that nest within the Saccharomycetales; the *Saccharomyces*-specific Mph2/3 clade is indicated. Circles denote branches with >90% bootstrap support. Colored bars outside the tree show genus-level taxonomic assignment, and the inset circular tree shows the Saccharomycotina species phylogeny (Shen et al. 2018) with those genera colored; *Zygo/torulaspora* represents *Zygosaccharomyces*, *Zygotorulaspora*, and *Torulaspora*. The rooted maximum-likelihood tree can be found in supplementary fig. S6, Supplementary Material online. Newick-formatted trees are available in supplementary data S2 and S3, Supplementary Material online. The full MFS phylogeny can be found in supplementary fig. S5, Supplementary Material online. Scale bars indicate the number of substitutions per site.

to the *Lachancea* Agt1-like proteins and (ii) more notably, a well-supported clade of Agts from *Brettanomyces anomalus* and *B. bruxellensis*. The *Brettanomyces* species are documented recipients of numerous horizontal gene transfer events, including for genes involved in the metabolism of sucrose, an Agt1 substrate (Stambuk et al. 2000; Woolfit et al. 2007; Roach and Borneman 2020). Notably, *B. bruxellensis* is commonly associated with brewing environments, where its propensity to vigorously consume diverse sugars and independent evolution of aerobic fermentation make it a frequent contaminant and occasional desired contributor (Rozpedowska et al. 2011; Colomer et al. 2019; Colomer et al. 2020).

Surprisingly, the clade containing high-specificity *Saccharomyces* maltose transporters only included taxa from closely related species in the genera *Saccharomyces* and *Lachancea*, as well as one protein each from *Zygotorulaspora florentina* and *Zygosaccharomyces kombuchaensis* (Fig. 6b). Among the high-specificity Agts, the

Mph2/3 clade was further restricted to *S. kudriavzevii*, *S. mikatae*, *S. paradoxus*, and *S. cerevisiae* (Fig. 6b), which is consistent with an origin in the common ancestor of these species following their split from *S. arboricola* and a recent segmental duplication in *S. cerevisiae* (*S. jurei* is absent from this dataset). The sister clade to the high-specificity proteins contained generalist Agts from *Saccharomyces*, *Torulaspora*, and *Zygotorulaspora* species, with deeper branches to *Kluyveromyces* and *Lachancea* homologs (Fig. 6b). We thus conclude that the high-specificity transporters typified by *S. cerevisiae* Mal31, including *S. eubayanus* MalT4 and MalT3, form a clade restricted to Saccharomycetales.

Generalist-Like Transporters Are Quantitatively Correlated With Growth on α-glucosides

Our phylogenetic analyses suggested that the high-specificity Agts are evolutionarily and functionally derived from a

generalist ancestor. In this model, the vast array of uncharacterized Agt clade proteins encoded by diverse yeast species should include generalist transporters or transporters that became subfunctionalized following duplication of a generalist ancestor, and their presence should support growth on substrates of the generalist Agts. We collected quantitative growth measurements for 287 of the 332 species in our phylogenetic dataset on three sugars that are substrates of the generalist transporter *S. cerevisiae* Agt1 but not of the high-specificity transporters: maltotriose, trehalose, and methyl-α-glucoside (Han et al. 1995; Stambuk et al. 1999; Stambuk and Araujo 2001; Alves et al. 2008; Brown et al. 2010). We found many species across the Saccharomycotina to be capable of vigorous growth on these sugars as a sole carbon source (Fig. 7a). Growth on all three α-glucosides was nearly ubiquitous among Serinales, a speciose order with a high incidence of carbon niche breadth generalists (Opulente et al. 2024). Most notably, growth on maltotriose was widespread across the yeast subphylum, in contrast to the documented rarity of this trait in the model genus *Saccharomyces* (Duval et al. 2010; Gallone et al. 2018; Langdon et al. 2020; Hutzler et al. 2021; Gyurchev et al. 2022; Peris et al. 2023). This metabolic deficiency was concomitant with the paucity of generalist-like Agt proteins encoded in Saccharomycetales genomes, which was similarly not representative of other yeast orders (Fig. 7b; $P = 1.9 \times 10^{-13}$). Indeed, patterns of α-glucoside growth qualitatively tracked the presence of genes encoding Agt proteins, with both subject to clear evolutionary shifts including losses (e.g. Saccharomycodales, Sporopachydermiales, and Trigonopsidales; *Saturnispora*, *Zygosaccharomyces*, *Eremothecium*, *Kazachstania*, *Nakaseomyces*, *Naumovozyma*, and *Tetrapisispora* spp.) and amplifications (*Debaryomyces*, *Metschnikowia*, and *Kuraishia* spp.; subclades of Phaffomycetales, Dipodascales, Pichiales, and Lipomycetales). We used phylogenetically corrected least squares regressions (PGLSs) to statistically test the strength of the correlation between Agt count and growth on each of the three tested Agt1 substrates (Fig. 7c). We detected significant positive correlations between Agt count and growth on each of the three α-glucosides ($P \leq 0.007$). Thus, the generalist-like Agts detected in most Saccharomycotina genomes are likely to be true generalist transporters or recently subfunctionalized derivatives.

Discussion

In the present work, we sought to understand how novel function could evolve in a model yeast α-glucoside transporter. To this end, we dissected the molecular basis of maltotriose transport in MalT434, which represents one of the most evolutionarily recent functional innovations in this family. We found that, in this chimeric protein, novel function is an emergent property of extensive additive and nonadditive interactions between multiple protein regions and multiple residues on TMHs 7, 10, and 11 (Figs. 2 to 4). We observed that even conservative amino acid changes, as well as residues not predicted to interact with the substrate, had significant and unexpected effects

on maltotriose transport (Figs. 3 and 5). We also found evidence that the stringency of side chain physicochemical requirements likely differs substantially between crucial residues (Fig. 5; supplementary fig. S4, Supplementary Material online). Taken together, these results demonstrate that the evolution of novel function in a high-specificity Agt is highly constrained, which is consistent with recent observations (Hatanaka et al. 2022). In this model, while gene conversion between homologs may not be strictly required for the evolution of novel function in this family, it may indeed be the only remotely probable way that all the necessary interacting residues can readily be assembled in a single molecule, even if paralogs are free to sample neutral or deleterious mutational steps.

The gene conversion events leading to novel function in high-specificity yeast Agts share striking parallelism at both the sequence and structural scales. For example, the portions of Mty1 inferred to derive from different parental proteins encompass many of the same regions that we identified as having crucial interactions in MalT434 (supplementary fig. S7a, Supplementary Material online). Even more strikingly, the homologous residues at five of the seven sites that affect maltotriose transport in MalT434 are conserved in Mty1 (supplementary fig. S7b, Supplementary Material online). At the other two sites, Mty1 possesses amino acids that support reduced, but significant, growth in MalT434 (C505S and T512I). While many of the same sites likely contribute to novel function in both of these recombinant transporters, specific amino acids at key sites are still likely context dependent, which makes functional evolution both more difficult to predict and to engineer in this family.

Compounding this difficulty is the cryptic nature of sites that we empirically determined to influence maltotriose transport, but which are unlikely to interact with the substrate (Fig. 1). These substitutions may affect subtle changes to the overall conformation of the transporter, especially where they have the potential to interact with other protein regions that are proximal in tertiary space (e.g. F468). Moreover, there is a growing appreciation that, in yeast monosaccharide sugar porters, the fine-scale environment around the substrate-binding site plays a surprisingly large role in sugar recognition and specificity, both by shaping an accommodating binding pocket and through interactions between substrate-interacting and noninteracting residues within Van der Waals distance (Kasahara et al. 2009; Drew et al. 2021).

In MalT434, more concrete hypotheses can be made about the molecular contributions of other sites important for novel substrate transport. Molecular docking analyses place the maltotriose ligand in close proximity to the key sites on TMH 7 and TMH 11 (supplementary fig. S8, Supplementary Material online), with several of the sugar hydroxyl groups capable of engaging in a hydrogen-bonding network with the side chains of polar amino acid residues at those sites. Of the substitutions in MalT434 that face the transport channel, all three have
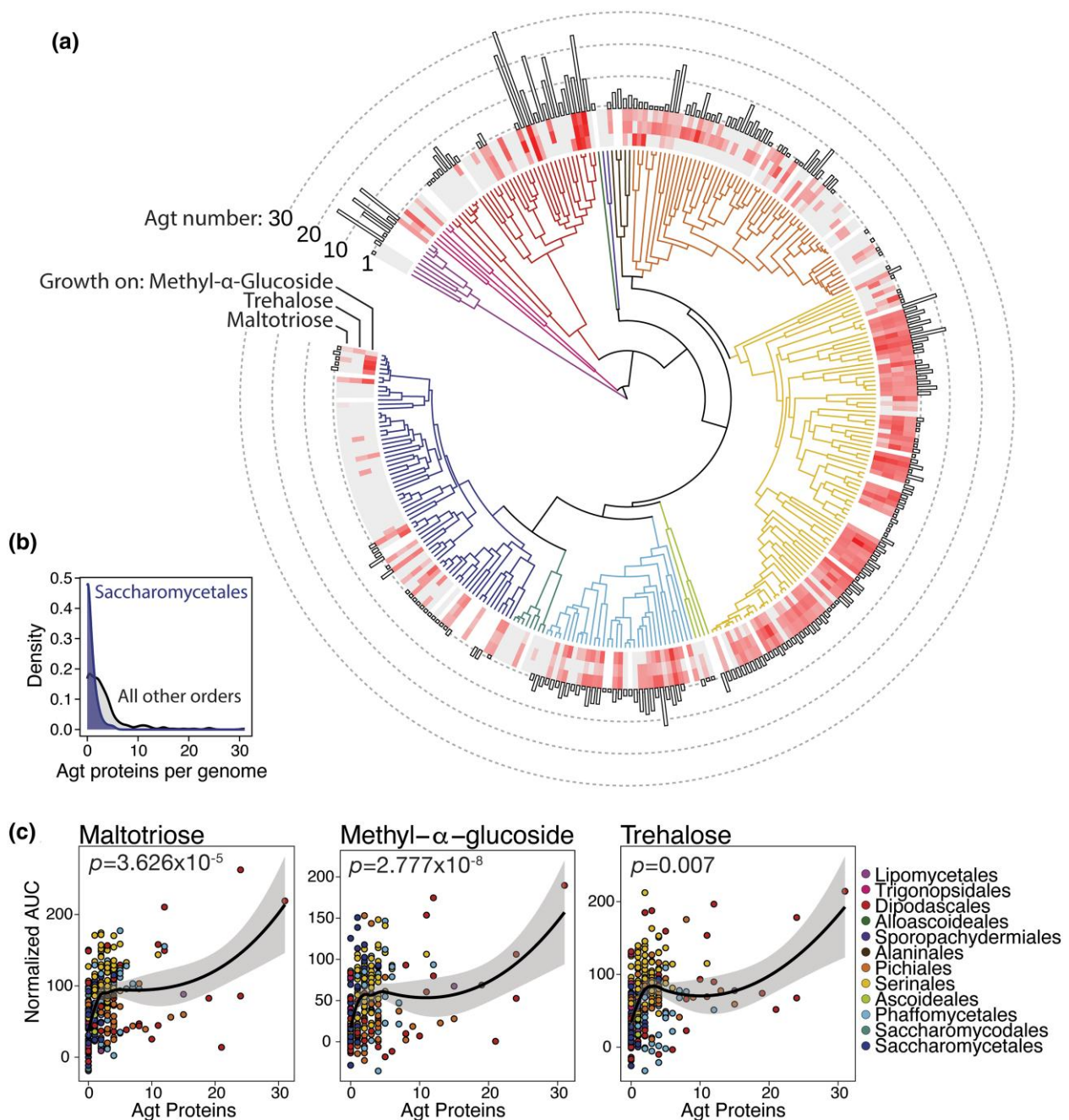
Fig. 7. Species with Agt proteins grow on Agt1-specific substrates. a) Time-calibrated phylogeny of 332 Saccharomycotina species (Shen et al. 2018) with branches colored [key in c)] by taxonomic order (Groenewald et al. 2023). Heatmaps around the tree show growth (normalized AUC) on α-glucosides: methyl-α-glucoside (inner ring), trehalose (middle ring), and maltotriose (outer ring). Gray boxes denote no growth above background; white boxes represent unsampled species. The bar chart shows the number of proteins in the α-glucoside transporter clade for each genome. b) Generalist Agt content of Saccharomycetales genomes is not representative. Density plots show distributions of the number of Agt clade proteins per genome for Saccharomycetales species and species from all other orders. c) Scatterplots of Agt clade transporter count vs. growth on each α-glucoside. Each species is represented by a point, colored by taxonomic order. Lines and shaded regions are LOESS-smoothed regressions of the untransformed data; inset P-values are from PGLSs.

polar and hydrogen bond–competent side chains of small-to-medium size; in wild-type MalT4, the residues at these sites have bulkier and/or hydrophobic side chains. Similarly, at the crucial 379 site on TMH 7, the permissive serine has a much smaller side chain than the prohibitive glutamine. Either of the prohibitive residues at 379 and the other crucial site 505 might introduce steric clashes

with the terminal glucopyranose moiety of maltotriose (supplementary fig. S8c, Supplementary Material online), even though they themselves are likely capable of hydrogen bonding with the substrate. Notably, the residue at position 379 may be involved in coupling substrate binding to gating during the transition to the occluded state (Drew et al. 2021), a key determinant of substrate

recognition that involves more tightly embedding the sugar molecule in its binding site within the transport channel. In wild-type MalT4, position 379 has the smaller serine residue, while sites along TMH 11 have bulkier amino acids; in wild-type MalT3, position 379 has the larger glutamine residue, but TMH 11 has smaller, hydrophilic residues. Thus, in each native maltose transporter, the steric constraint of the transport channel may be finely tuned at coevolving sites along TMH 7 and TMH 11 to accommodate maltose with higher affinity and specificity, which occur at the expense of steric exclusion of other substrates, such as maltotriose (supplementary fig. S8e, Supplementary Material online). This model is consistent with the crucial role of amino acid side chain length in shaping substrate specificity in some monosaccharide sugar porters (Kasahara et al. 2011; Drew et al. 2021), notwithstanding that we also detected a complex interaction between size and biochemical properties at the key 505 site.

The difficulty of functional innovation in the high-specificity Agts begs the question of how the related generalist Agts are capable of transporting not only maltose and maltotriose but also a diverse range of substrates. If the generalist transporters had evolved from a more specific ancestor, as has been suggested (Pougach et al. 2014), their extant substrate range would imply multiple bouts of highly constrained functional evolution. To determine when and how this broad substrate specificity may have evolved in the generalist Agts, we reconstructed the yeast sugar porter phylogeny from 332 newly annotated, representative Saccharomycotina genomes encompassing more than 400 My of evolution (supplementary fig. S5, Supplementary Material online). This analysis showed that, somewhat unexpectedly, the high-specificity Agts are a derived clade within the generalist-like Agts (Fig. 6a). The copy number of these putative generalist Agts encoded by yeast genomes is strongly predictive of growth on Agt1-exclusive substrates (Fig. 7), which further supports the conclusion that these proteins are likely bona fide generalists. The evolution of maltotriose transport by high-specificity Agts is thus better regarded as a reacquisition of ancestral function than the de novo evolution of a truly novel function within this protein family.

It remains subject to debate whether the general trend of protein evolution is directional: from less to more intrinsically specific (Bridgham et al. 2006; Khersonsky and Tawfik 2010; Copley 2012; Steindel et al. 2016; Wheeler et al. 2016; Wheeler and Harms 2021). Multiple lines of evidence now suggest that this mode predominates among genes involved in α-glucoside metabolism in yeasts. In addition to the α-glucoside transporters, both the α-glucosidases of *S. cerevisiae* and the transcriptional activators that regulate the structural metabolic genes likely evolved from promiscuous ancestral proteins that optimized subfunctions following duplication events, rendering them specific for different α-glucosides (Brown et al. 2010; Voordeckers et al. 2012; Pougach et al. 2014). The extent of intramolecular epistasis apparent in the high-specificity Agts, which may arise both from intraprotein and protein–substrate interactions, may provide an explanation for the inherent difficulty of re-evolving

maltotriose transport in these proteins. Functional entrenchment by historical contingency and epistasis is well documented, and the irreversibility of evolutionary trajectories at the molecular level may be a widespread phenomenon (Ortlund et al. 2007; Bridgham et al. 2009; Soylemez and Kondrashov 2012; Harms and Thornton 2014; Bank et al. 2015; Podgornaia and Laub 2015; Shah et al. 2015; Starr and Thornton 2016; Starr et al. 2017; Starr et al. 2018; Ben-David et al. 2020; Xie et al. 2021; Park et al. 2022). Although not directly tested here, there may be inherent tradeoffs between specificity and substrate affinity in yeast Agts (Stambuk and Araujo 2001; Salema-Oom et al. 2005; Hatanaka et al. 2022), which would suggest that walking back through the accumulated mutations that led to higher specificity in the Mal31-like transporters would be likely to incur an immediate functional tradeoff and therefore fitness cost. The recurrent gene conversion events that enable maltotriose transport among members of this family may, therefore, represent the only meaningfully accessible route to bypass these deleterious intermediates, but the high degree of context dependence for mutational effects makes the prediction or engineering of this novel function difficult (Hatanaka et al. 2022).

Might the evolution of yeast sugar porters more broadly be organized along an axis of increasing specialization and specificity? This family encompasses functionally diverse transporters with varying specificities for different mono- and di-saccharides and sugar alcohols; notably, functionally similar proteins are not monophyletic across the family (Donzella et al. 2023). Our phylogenetic analysis of these proteins places the Agts, which may retain some glucose transport capacity (Wieczorke et al. 1999), as a deeply branching sister clade to most of the broader family (supplementary fig. S5, Supplementary Material online). These results imply multiple bouts of functional specialization from a highly promiscuous ancestor, in some cases starting from partially subfunctionalized ancestral proteins, with the Agts perhaps remaining the most representative of the ancestral multifunctionality. Supporting this notion, the Agt gene family is a key contributor to the carbon niche breadth of more than 1,000 yeasts (David et al. 2024). While the extant diversity of yeast sugar porters has generally been regarded as an example of functional diversification (i.e. highly plastic gains of novel substrate affinity; [Brown et al. 2010; Hatanaka et al. 2022; Donzella et al. 2023]), the evolution of this important gene family may have followed a very different mode. In the former model, functional diversification by neofunctionalization follows duplication of ancestral transporter genes, whereas our analyses suggest that duplications in this gene family may be primarily followed by subfunctionalizing escapes from adaptive conflict (Hughes 1994; Hittinger and Carroll 2007; Des Marais and Rausher 2008), wherein transporters can gain increased specificity and affinity for a narrow substrate range at the expense of other ancestral ligands.

These two models have distinct implications for the myriad biotechnological applications predicated upon sugar consumption by yeasts, which might be targets for improvement by protein engineering. If extant transporters

are indeed highly plastic and evolvable, shifting or expanding their substrate range should be relatively simple. If, on the other hand, they have undergone entrenched specialization, they may be inherently less evolvable (Bridgham et al. 2009; Starr et al. 2018; Wheeler and Harms 2021). Results here and elsewhere (Hatanaka et al. 2022) support the latter corollary. However, this model also implies that reconstructed ancestral proteins, or even generalist extant proteins from this clade, might both possess desirable properties and be inherently highly amenable to engineering, mutagenesis, or directed evolution approaches.

Materials and Methods

Strains and Cultivation Conditions

Saccharomyces eubayanus strains, plasmids, and oligonucleotides used in this work are listed in supplementary tables S1 and S2, Supplementary Material online. Yeasts were propagated on yeast extract, peptone, and dextrose (YPD) medium (1% yeast extract, 2% peptone, and 2% glucose) supplemented with 400 mg/L G418 and/or 50 mg/L nourseothricin (CloNAT) as appropriate and cryopreserved in 15% glycerol at −80 °C for long-term storage.

Transformation of *S. eubayanus* was performed by the polyethylene glycol/lithium acetate/carrier DNA method (Gietz and Schiestl 2007) with minor modifications (Baker and Hittinger 2019). Clustered regularly interspaced short palindromic repeat-Cas9 (CRISPR-Cas9)-mediated gene deletions and insertions were achieved by cotransformation of pXIPHOS vectors (Kuang et al. 2018) and repair templates for homologous recombination. Repair templates were purified polymerase chain reaction (PCR) products consisting of single linear fragments, multiple linear fragments for in vivo assembly, or recombinant amplicons generated by overlap extension PCR, depending on the application. All repair templates were amplified using Phusion polymerase (New England Biolabs) per the manufacturer's instructions and purified using QiaQuick or MinElute spin columns (Qiagen).

We assessed transporter function via expression from the native *MALT4* locus in yHJC207, a haploid derivative of the wild strain yHKS210 that was constructed as previously described (Crandall et al. 2023). Because the *MALT2* and *MALT4* loci are recent duplicates and almost identical at the nucleotide level, transporter variants were inserted into both loci out of necessity. Both *MALT2* and *MALT4* were simultaneously deleted using CRISPR-Cas9 and replaced with *kanMX*. Novel transporter variants, as well as *MALT434* and *S. eubayanus AGT1*-positive controls, were subsequently inserted into both loci by cotransformation with a pXIPHOS vector expressing Cas9 and a gRNA targeting *kanMX* (Lee et al. 2021). Transformants were cured of plasmids, and the inserted alleles were sequenced.

Quantitative Growth Measurements of *S. eubayanus* Strains

Strains were streaked to single colonies on YPD plates, arrayed in 96-well plates in a randomized layout, and

preultured in YPD at room temperature for 72 h with gentle shaking. Precultures were serially diluted in minimal medium (0.5% ammonium sulfate and 0.017% Yeast Nitrogen Base) and inoculated into minimal medium containing 2% sugars in 96-well plates at a final dilution of 10^{-4}. OD_{600} was measured every hour for 7 d using a SPECTROstar Omega plate reader (BMG Labtech) equipped with a microplate stacker. Raw growth data were summarized using GCAT (Bukhman et al. 2015). Area under the curve (AUC) measurements for growth on maltotriose, normalized to a common negative control within each experiment, were used as a response variable in linear models with protein identity (MalT3 or MalT4) at each domain or at key amino acid sites as categorical predictor variables. The effects of protein identity at some single regions and for many pairwise interactions could not be estimated due to singularities. We tested for evidence of epistasis by statistically comparing additive models and those with interaction terms (Li and Fay 2019). The amino acid properties compiled to test associations with transporter function included chemical composition, polarity, and volume (Grantham 1974), aromaticity (Xia and Li 1998), hydropathy (JANIN 1979; Kyte and Doolittle 1982; Hopp and Woods 1983; Eisenberg et al. 1984; Rose et al. 1985; Engelman et al. 1986; Cornette et al. 1987), and BLOSUM similarity (Henikoff and Henikoff 1992). Some matrices were compiled from Braun (2018). For dimensionality reduction, BLOSUM similarity was omitted.

Quantitative Growth Measurements of Saccharomycotina Yeasts

Growth on α-glucosides was measured for the strains whose genome annotations were analyzed, which were primarily the type strains for their respective species. Strain information, including taxonomic order (Groenewald et al. 2023), major clade (Shen et al. 2018), and updated annotation mapping, can be found in supplementary table S3, Supplementary Material online. Cryopreserved strains were inoculated directly to YPD in 96-well plates and incubated for 7 d at room temperature. Some slow-growing species failed to revive during this time frame and were removed from further analysis, and we did not phenotype opportunistic pathogens, ultimately resulting in data for 287 species. Precultures were inoculated to minimal medium with 1% sugar or no added carbon source using a pinning tool, incubated for 7 d at room temperature, and re-inoculated to new plates containing the same medium. OD_{600} of the second round of growth was measured every hour using a SPECTROstar Omega plate reader (BMG Labtech) equipped with a microplate stacker. The growth experiments were performed four times independently. Raw growth data were summarized using Growthcurver (Sprouffske and Wagner 2016). Wells with poor model fits were discarded, and each curve was manually inspected to identify species with unreliable growth curves (Opulente et al. 2024). Growth on each carbon source was normalized to the average growth of the same species in medium with no added carbon to control for background growth.

Caper (cran.r-project.org/web/packages/caper/index.html) was used to fit PGLSs to growth data and square root–transformed Agt number, using the rooted maximum likelihood species phylogeny (Shen et al. 2018).

Structure Prediction and Analyses

Structural models for MalT434 were generated using four different softwares: AlphaFold2 (Jumper et al. 2021), Phyre2 (Kelley et al. 2015), I-TASSER (Yang et al. 2015), and SWISS-MODEL (Waterhouse et al. 2018a). All gave extremely similar results across the structured region (mean and standard deviation pairwise RMSD: 1.61 ± 0.51Å), and AlphaFold2 models for all proteins of interest were generated and used for further analysis. Docking of maltotriose was performed using SwissDock (Grosdidier et al. 2011). Structure models and docking results were visualized in PyMol v2.5 (Schrödinger, LLC).

Genome Annotation

To improve the quality of existing gene models, publicly available genome assemblies of 332 Saccharomycotina yeast species (Shen et al. 2018) were re-annotated de novo. For consistency, we retained the assembly and species names, although some species have since been renamed; consult MycoBank (www.mycobank.org) for the most up-to-date taxonomic information. Repetitive sequences were softmasked with RepeatMasker v4.1.2, and protein-coding genes were annotated using ab inito predictors AUGUSTUS v3.4.0 (Stanke et al. 2008) and GeneMark-EP+ v4.6.1 (Brůna et al. 2020) in BRAKER (Brůna et al. 2021), with Saccharomycetes proteins in OrthoDB v10 (Kriventseva et al. 2019) as homology evidence and using the –fungus mode. Where applicable, the longest transcript of each gene was retained. BUSCO v5.7.0 (Manni et al. 2021) was used to assess the completeness of the new and preexisting genome annotations using single-copy yeast orthologs in OrthoDB v10 (Waterhouse et al. 2018b).

This approach was chosen so as to generate a useful community resource in two ways: first, to enable direct comparisons with a larger, partially overlapping dataset of yeast genomes published recently (Opulente et al. 2024), which were annotated using identical methods, and second, to facilitate future studies by significantly improving the quality of annotations for the widely used 332-genome dataset. Median annotation completeness was increased from 94.6% to 98.8%, while the median percentage of missing BUSCO genes decreased to 0.9% from 4.6% (both $P < 2.2 \times 10^{-16}$, two-sided t-tests; supplementary fig. S10, Supplementary Material online). Supplementary table S4, Supplementary Material online documents BUSCO analyses of existing and updated annotations for all genomes. The full updated annotations in protein and nucleotide FASTA, GFF3, and GTF formats are available on figshare: https://doi.org/10.6084/m9.figshare.26166640.

Phylogenetic Analyses

The amino acid translations of the newly predicted protein-coding genes were queried by BLASTp+ v2.9

(Camacho et al. 2009) using characterized S. cerevisiae sugar transporters (Mal31, Agt1, Gal2, Hxt1-5, and Hxt7) retrieved from Saccharomyces Genome Database (SGD) (Wong et al. 2023). BLAST subjects <400 or >1,000 amino acids in length were discarded to remove partial or fused annotations, based on distributions of sugar porter length in TCDB (Saier et al. 2006, 2021). Remaining proteins were annotated with their most similar S. cerevisiae homolog using a reciprocal BLASTp search against all translated open reading frames in S. cerevisiae, which were retrieved from SGD. Protein sequences were aligned using the E-INS-i strategy of MAFFT v7.222 (Katoh et al. 2002, 2005; Katoh and Standley 2013), and the alignment was trimmed with trimAL v1.4.22 (Capella-Gutiérrez et al. 2009) using the –gappyout parameter. The phylogeny was inferred using IQ-TREE v2.2.2.7 (Minh et al. 2020) with 1,000 bootstraps (Hoang et al. 2018) and automatic substitution model selection (Kalyaanamoorthy et al. 2017). Due to the high conservation of MFS proteins, this dataset contained a small proportion of MFS proteins that do not belong to the sugar porter family, primarily belonging to the drug:proton antiporter family. These were retained in the alignment and tree inference to test the assumption of sugar porter monophyly. As expected, the sugar porters and other MFS proteins formed well-supported reciprocally monophyletic clades. The α-glucoside transporter phylogeny was refined by re-aligning the proteins from that clade and inferring the phylogeny as before, albeit with 10 independent runs of IQ-TREE with 10,000 bootstrap replicates each and secondary branch support assessment by Shimodaira-Hasegawa-like approximate likelihood ratio test. Trees were visualized and annotated in iTOL (Letunic and Bork 2021).

Supplementary Material

Supplementary material is available at Molecular Biology and Evolution online.

· **MBE**

Funding

Conflict of Interest

The Wisconsin Alumni Research Foundation has filed a provisional patent application on the technologies described herein with J.G.C. and C.T.H. as inventors. Strains are available for noncommercial, academic use under a material transfer agreement. A.R. is a scientific consultant for LifeMine Therapeutics, Inc.

Data Availability

New genome annotations for Saccharomycotina species are available on figshare: https://doi.org/10.6084/m9.figshare.26166640.

References

Abramson J, Smirnova I, Kasho V, Verner G, Kaback HR, Iwata S. Structure and mechanism of the lactose permease of *Escherichia coli*. Science. 2003:**301**(5633):610–615. https://doi.org/10.1126/science.1088196.

Alves SL, Herberts RA, Hollatz C, Trichez D, Miletti LC, de Araujo PS, Stambuk BU. 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):1494–1501. https://doi.org/10.1128/AEM.02570-07.

Arguello JR, Chen Y, Yang S, Wang W, Long M. Origination of an X-linked testes chimeric gene by illegitimate recombination in *Drosophila*. PLoS Genet. 2006:**2**(5):e77. https://doi.org/10.1371/journal.pgen.0020077.

Baker EP, Hittinger CT. Evolution of a novel chimeric maltotriose transporter in *Saccharomyces eubayanus* from parent proteins unable to perform this function. PLoS Genet. 2019:**15**(4):e1007786. https://doi.org/10.1371/journal.pgen.1007786.

Baker EP, Sayegh R, Kohler KM, Borman W, Goodfellow CK, Brush ER, Barber MF. Evolution of host-microbe cell adherence by receptor domain shuffling. Elife. 2022:**11**:e73330. https://doi.org/10.7554/eLife.73330.

Bank C, Hietpas RT, Jensen JD, Bolon DNA. A systematic survey of an intragenic epistatic landscape. Mol Biol Evol. 2015:**32**(1):229–238. https://doi.org/10.1093/molbev/msu301.

Ben-David M, Soskine M, Dubovetskyi A, Cherukuri K-P, Dym O, Sussman JL, Liao Q, Szeler K, Kamerlin SCL, Tawfik DS. Enzyme evolution: an epistatic ratchet versus a smooth reversible transition. Mol Biol Evol. 2020:**37**(4):1133–1147. https://doi.org/10.1093/molbev/msz298.

Bergin SA, Allen S, Hession C, Cinnéide EÓ, Ryan A, Byrne KP, Cróinín TÓ, Wolfe KH, Butler G, Morrissey J. Identification of European isolates of the lager yeast parent *Saccharomyces eubayanus*. FEMS Yeast Res. 2022:**22**(1):1–9. https://doi.org/10.1093/femsyr/foac053.

Bittihn P, Tsimring LS. Gene conversion facilitates adaptive evolution on rugged fitness landscapes. Genetics. 2017:**207**(4):1577–1589. https://doi.org/10.1534/genetics.117.300350.

Bosshart PD, Fotiadis D. Secondary active transporters. Subcell Biochem. 2019:**92**:275–299. https://doi.org/10.1007/978-3-030-18768-2_9.

Braun EL. An evolutionary model motivated by physicochemical properties of amino acids reveals variation among proteins. Bioinformatics. 2018:**34**(13):i350–i356. https://doi.org/10.1093/bioinformatics/bty261.

Brickwedde A, Brouwers N, van den Broek M, Gallego Murillo JS, Fraiture JL, Pronk JT, Daran JG. Structural, physiological and regulatory analysis of maltose transporter genes in *Saccharomyces eubayanus* CBS 12357^T. Front Microbiol. 2018:**9**:1786. https://doi.org/10.3389/fmicb.2018.01786.

Bridgham JT, Carroll SM, Thornton JW. Evolution of hormone-receptor complexity by molecular exploitation. Science. 2006:**312**(5770):97–101. https://doi.org/10.1126/science.1123348.

Bridgham JT, Ortlund EA, Thornton JW. An epistatic ratchet constrains the direction of glucocorticoid receptor evolution. Nature. 2009:**461**(7263):515–519. https://doi.org/10.1038/nature08249.

Brouwers N, Brickwedde A, Gorter de Vries AR, van den Broek M, Weening SM, van den Eijnden L, Diderich JA, Bai FY, Pronk JT, Daran JMG. Himalayan *Saccharomyces eubayanus* genome sequences reveal genetic markers explaining heterotic maltotriose consumption by *Saccharomyces pastorianus* hybrids. Appl Environ Microbiol. 2019a:**85**(22):e01516–e01519. https://doi.org/10.1128/AEM.01516-19.

Brouwers N, Gorter de Vries AR, van den Broek M, Weening SM, Elink Schuurman TD, Kuijpers NGA, Pronk JT, Daran J-MG. In vivo recombination of *Saccharomyces eubayanus* maltose-transporter genes yields a chimeric transporter that enables maltotriose fermentation. PLoS Genet. 2019b:**15**(4):e1007853. https://doi.org/10.1371/journal.pgen.1007853.

Brown CA, Murray AW, Verstrepen KJ. Rapid expansion and functional divergence of subtelomeric gene families in yeasts. Curr Biol. 2010:**20**(10):895–903. https://doi.org/10.1016/j.cub.2010.04.027.

Brůna T, Hoff KJ, Lomsadze A, Stanke M, Borodovsky M. BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. NAR Genom Bioinform. 2021:**3**:lqaa108. doi: 10.1093/nargab/lqaa108.

Brůna T, Lomsadze A, Borodovsky M. GeneMark-EP+: eukaryotic gene prediction with self-training in the space of genes and proteins. NAR Genom Bioinform. 2020:**2**(2):lqaa026. https://doi.org/10.1093/nargab/lqaa026.

Bukhman YV, DiPiazza NW, Piotrowski J, Shao J, Halstead AGW, Bui MD, Xie E, Sato TK. Modeling microbial growth curves with GCAT. Bioenergy Res. 2015:**8**(3):1022–1030. https://doi.org/10.1007/s12155-015-9584-3.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. BMC Bioinformatics. 2009:**10**(1):421. https://doi.org/10.1186/1471-2105-10-421.

Campbell E, Kaltenbach M, Correy GJ, Carr PD, Porebski BT, Livingstone EK, Afriat-Jurnou L, Buckle AM, Weik M, Hollfelder F, et al. The role of protein dynamics in the evolution of new enzyme function. Nat Chem Biol. 2016:**12**(11):944–950. https://doi.org/10.1038/nchembio.2175.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. Trimal: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009:**25**(15):1972–1973. https://doi.org/10.1093/bioinformatics/btp348.

Casa-Villegas M, Polaina J, Marín-Navarro J. Cellobiose fermentation by *Saccharomyces cerevisiae*: comparative analysis of intra versus extracellular sugar hydrolysis. Process Biochem. 2018:**75**:59–67. https://doi.org/10.1016/j.procbio.2018.09.005.

Cheng CHC. Evolution of the diverse antifreeze proteins. Curr Opin Genet Dev. 1998:**8**(6):715–720. https://doi.org/10.1016/S0959-437X(98)80042-7.

Cheng Q, Michels CA. *MAL11* and *MAL61* encode the inducible high-affinity maltose transporter of *Saccharomyces cerevisiae*. J Bacteriol. 1991:**173**(5):1817–1820. https://doi.org/10.1128/jb.173.5.1817-1820.1991.

Christin P-A, Salamin N, Savolainen V, Duvall MR, Besnard G. C4 photosynthesis evolved in grasses via parallel adaptive genetic changes. Curr Biol. 2007:**17**(14):1241–1247. https://doi.org/10.1016/j.cub.2007.06.036.

Ciccarelli FD, von Mering C, Suyama M, Harrington ED, Izaurralde E, Bork P. Complex genomic rearrangements lead to novel primate gene function. Genome Res. 2005:**15**(3):343–351. https://doi.org/10.1101/gr.3266405.

Clark AG, Glanowski S, Nielsen R, Thomas PD, Kejariwal A, Todd MA, Tanenbaum DM, Civello D, Lu F, Murphy B, et al. Inferring non-neutral evolution from human-chimp-mouse orthologous gene trios. Science. 2003:**302**(5652):1960–1963. https://doi.org/10.1126/science.1088821.

Colomer MS, Chailyan A, Fennessy RT, Olsson KF, Johnsen L, Solodovnikova N, Forster J. Assessing population diversity of *Brettanomyces* yeast species and identification of strains for brewing applications. Front Microbiol. 2020:**11**:637. https://doi.org/10.3389/fmicb.2020.00637

Colomer MS, Funch B, Forster J. The raise of *Brettanomyces* yeast species for beer production. Curr Opin Biotechnol. 2019:**56**:30–35. https://doi.org/10.1016/j.copbio.2018.07.009.

Copley SD. Toward a systems biology perspective on enzyme evolution. J Biol Chem. 2012:**287**(1):3–10. https://doi.org/10.1074/jbc.R111.254714.

Cornette JL, Cease KB, Margalit H, Spouge JL, Berzofsky JA, DeLisi C. Hydrophobicity scales and computational techniques for detecting amphipathic structures in proteins. J Mol Biol. 1987:**195**(3):659–685. https://doi.org/10.1016/0022-2836(87)90189-6.

Cousseau FEM, Alves SL, Trichez D, Stambuk BU. Characterization of maltotriose transporters from the *Saccharomyces eubayanus* subgenome of the hybrid *Saccharomyces pastorianus* lager brewing yeast strain Weihenstephan 34/70. Lett Appl Microbiol. 2013:**56**(1):21–29. https://doi.org/10.1111/lam.12011.

Crandall JG, Fisher KJ, Sato TK, Hittinger CT. Ploidy evolution in a wild yeast is linked to an interaction between cell type and metabolism. PLoS Biol. 2023:**21**(11):e3001909. https://doi.org/10.1371/journal.pbio.3001909.

Cubillos FA, Gibson B, Grijalva-Vallejos N, Krogerus K, Nikulin J. Bioprospecting for brewers: exploiting natural diversity for naturally diverse beers. Yeast. 2019:**36**(6):383–398. https://doi.org/10.1002/yea.3380.

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):809–814. https://doi.org/10.1073/pnas.022240299.

David KT, Schraiber JG, Crandall JG, Labella AL, Opulente DA, Harrison M-C, Wolters JF, Zhou X, Shen X-X, Groenewald M, et al. Convergent expansions of keystone gene families drive metabolic innovation in a major eukaryotic clade. bioRxiv 604484. https://doi.org/10.1101/2024.07.22.604484, 23 July 2024, preprint: not peer reviewed.

Day RE, Higgins VJ, Rogers PJ, Dawes IW. Characterization of the putative maltose transporters encoded by YDL247w and YJR160c. Yeast. 2002:**19**(12):1015–1027. https://doi.org/10.1002/yea.894.

Deng D, Sun P, Yan C, Ke M, Jiang X, Xiong L, Ren W, Hirata K, Yamamoto M, Fan S, et al. Molecular basis of ligand recognition and transport by glucose transporters. Nature. 2015:**526**(7573):391–396. https://doi.org/10.1038/nature14655.

Deng D, Xu C, Sun P, Wu J, Yan C, Hu M, Yan N. Crystal structure of the human glucose transporter GLUT1. Nature. 2014:**510**(7503):121–125. https://doi.org/10.1038/nature13306

de Ruijter JC, Igarashi K, Penttilä M. The *Lipomyces starkeyi* gene *Ls120451* encodes a cellobiose transporter that enables cellobiose fermentation in *Saccharomyces cerevisiae*. FEMS Yeast Res. 2020:**20**(3):foaa019. https://doi.org/10.1093/femsyr/foaa019.

Des Marais DL, Rausher MD. Escape from adaptive conflict after duplication in an anthocyanin pathway gene. Nature. 2008:**454**(7205):762–765. https://doi.org/10.1038/nature07092

De Visser JAGM, Krug J. Empirical fitness landscapes and the predictability of evolution. Nat Rev Genet. 2014:**15**(7):480–490. https://doi.org/10.1038/nrg3744.

Dietvorst J, Londesborough J, Steensma HY. Maltotriose utilization in lager yeast strains: *MTT1* encodes a maltotriose transporter. Yeast. 2005:**22**(10):775–788. https://doi.org/10.1002/yea.1279.

Donzella L, Sousa MJ, Morrissey JP. Evolution and functional diversification of yeast sugar transporters. Essays Biochem. 2023:**67**(5):811–827. https://doi.org/10.1042/EBC20220233.

Dorus S, Vallender EJ, Evans PD, Anderson JR, Gilbert SL, Mahowald M, Wyckoff GJ, Malcom CM, Lahn BT. Accelerated evolution of nervous system genes in the origin of *Homo sapiens*. Cell. 2004:**119**(7):1027–1040. https://doi.org/10.1016/j.cell.2004.11.040.

Drew D, North RA, Nagarathinam K, Tanabe M. Structures and general transport mechanisms by the major facilitator superfamily (MFS). Chem Rev. 2021:**121**(9):5289–5335. https://doi.org/10.1021/acs.chemrev.0c00983.

Duval EH, Alves SL, Dunn B, Sherlock G, Stambuk BU. Microarray karyotyping of maltose-fermenting *Saccharomyces* yeasts with differing maltotriose utilization profiles reveals copy number variation in genes involved in maltose and maltotriose utilization. J Appl Microbiol. 2010:**109**(1):248–259. https://doi.org/10.1111/j.1365-2672.2009.04656.x.

Eisenberg D, Schwarz E, Komaromy M, Wall R. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. J Mol Biol. 1984:**179**(1):125–142. https://doi.org/10.1016/0022-2836(84)90309-7.

Engelman DM, Steitz TA, Goldman A. Identifying nonpolar transbilayer helices in amino acid sequences of membrane proteins. Annu Rev Biophys Biophys Chem. 1986:**15**(1):321–353. https://doi.org/10.1146/annurev.bb.15.060186.001541.

Fairhead C, Dujon B. Structure of *Kluyveromyces lactis* subtelomeres: duplications and gene content. FEMS Yeast Res. 2006:**6**(3):428–441. https://doi.org/10.1111/j.1567-1364.2006.00033.x.

Farwick A, Bruder S, Schadeweg V, Oreb M, Boles E. Engineering of yeast hexose transporters to transport D-xylose without inhibition by D-glucose. Proc Natl Acad Sci U S A. 2014:**111**(14):5159–5164. https://doi.org/10.1073/pnas.1323464111.

Gallone B, Mertens S, Gordon JL, Maere S, Verstrepen KJ, Steensels J. Origins, evolution, domestication and diversity of *Saccharomyces* beer yeasts. Curr Opin Biotechnol. 2018:**49**:148–155. https://doi.org/10.1016/j.copbio.2017.08.005.

Gibson B, Geertman J-MA, Hittinger CT, Krogerus K, Libkind D, Louis EJ, Magalhães F, Sampaio JP. New yeasts—new brews: modern

approaches to brewing yeast design and development. FEMS Yeast Res. 2017:**17**(4):1–13. https://doi.org/10.1093/femsyr/fox038.

Gietz RD, Schiestl RH. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc. 2007:**2**(1):31–34. https://doi.org/10.1038/nprot.2007.13.

Gong LI, Suchard MA, Bloom JD. Stability-mediated epistasis constrains the evolution of an influenza protein. Elife. 2013:**2**: e00631. https://doi.org/10.7554/eLife.00631.

Gordon JL, Byrne KP, Wolfe KH. Additions, losses, and rearrangements on the evolutionary route from a reconstructed ancestor to the modern *Saccharomyces cerevisiae* genome. PLoS Genet. 2009:**5**(5):e1000485. https://doi.org/10.1371/journal.pgen.1000485.

Grantham R. Amino acid difference formula to help explain protein evolution. Science. 1974:**185**(4154):862–864. https://doi.org/10.1126/science.185.4154.862.

Groenewald M, Hittinger CT, Bensch K, Opulente DA, Shen XX, Li Y, Liu C, LaBella AL, Zhou X, Limtong S, et al. A genome-informed higher rank classification of the biotechnologically important fungal subphylum *Saccharomycotina*. Stud Mycol. 2023:**105**(1): 1–22. https://doi.org/10.3114/sim.2023.105.01.

Grosdidier A, Zoete V, Michielin O. SwissDock, a protein-small molecule docking web service based on EADock DSS. Nucleic Acids Res. 2011:**39**(suppl):W270–W277. https://doi.org/10.1093/nar/gkr366.

Guan L, Hariharan P. X-ray crystallography reveals molecular recognition mechanism for sugar binding in a melibiose transporter MelB. Commun Biol. 2021:**4**(1):931. https://doi.org/10.1038/s42003-021-02462-x.

Guan L, Kaback HR. Lessons from lactose permease. Annu Rev Biophys Biomol Struct. 2006:**35**(1):67–91. https://doi.org/10.1146/annurev.biophys.35.040405.102005.

Guan L, Mirza O, Verner G, Iwata S, Kaback HR. Structural determination of wild-type lactose permease. Proc Natl Acad Sci U S A. 2007:**104**(39):15294–15298. https://doi.org/10.1073/pnas.0707688104.

Gyurchev NY, Coral-Medina Á, Weening SM, Almayouf S, Kuijpers NGA, Nevoigt E, Louis EJ. Beyond *Saccharomyces pastorianus* for modern lager brews: exploring non-*cerevisiae* *Saccharomyces* hybrids with heterotic maltotriose consumption and novel aroma profile. Front Microbiol. 2022:**13**:1025132. https://doi.org/10.3389/fmicb.2022.1025132.

Ha S-J, Kim H, Lin Y, Jang M-U, Galazka JM, Kim T-J, Cate JHD, Jin Y-S. Single amino acid substitutions in HXT2.4 from *Scheffersomyces stipitis* lead to improved cellobiose fermentation by engineered *Saccharomyces cerevisiae*. Appl Environ Microbiol. 2013:**79**(5): 1500–1507. https://doi.org/10.1128/AEM.03253-12.

Han E-K, Cotty F, Sottas C, Jiang H, Michels CA. Characterization of *AGT1* encoding a general alpha-glucoside transporter from *Saccharomyces*. Mol Microbiol. 1995:**17**(6):1093–1107. https://doi.org/10.1111/j.1365-2958.1995.mmi_17061093.x.

Hansen TF, Carter AJ, Chiu CH. Gene conversion may aid adaptive peak shifts. J Theor Biol. 2000:**207**(4):495–511. https://doi.org/10.1006/jtbi.2000.2189.

Hara KY, Kobayashi J, Yamada R, Sasaki D, Kuriya Y, Hirono-Hara Y, Ishii J, Araki M, Kondo A. Transporter engineering in biomass utilization by yeast. FEMS Yeast Res. 2017:**17**(7):61. https://doi.org/10.1093/femsyr/fox061.

Harms MJ, Thornton JW. Historical contingency and its biophysical basis in glucocorticoid receptor evolution. Nature. 2014:**512**(7513): 203–207. https://doi.org/10.1038/nature13410.

Hatanaka H, Toyonaga H, Ishida Y, Mizohata E, Ono E. Functional diversity and plasticity in the sugar preferences of *Saccharomyces* MALT transporters in domesticated yeasts. FEMS Yeast Res. 2022:**22**(1):1–10. https://doi.org/10.1093/femsyr/foac055.

Henderson R, Poolman B. Proton-solute coupling mechanism of the maltose transporter from *Saccharomyces cerevisiae*. Sci Rep. 2017:**7**(1):14375. https://doi.org/10.1038/s41598-017-14438-1.

Henikoff S, Henikoff JG. Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci U S A. 1992:**89**(22):10915–10919. https://doi.org/10.1073/pnas.89.22.10915.

Hittinger CT, Carroll SB. Gene duplication and the adaptive evolution of a classic genetic switch. Nature. 2007:**449**(7163): 677–681. https://doi.org/10.1038/nature06151.

Hittinger CT, Steele JL, Ryder DS. Diverse yeasts for diverse fermented beverages and foods. Curr Opin Biotechnol. 2018:**49**: 199–206. https://doi.org/10.1016/j.copbio.2017.10.004.

Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol. 2018:**35**(2):518–522. https://doi.org/10.1093/molbev/msx281.

Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, Crossland JP. A single amino acid mutation contributes to adaptive beach mouse color pattern. Science. 2006:**313**(5783):101–104. https://doi.org/10.1126/science.1126121.

Hopp TP, Woods KR. A computer program for predicting protein antigenic determinants. Mol Immunol. 1983:**20**(4):483–489. https://doi.org/10.1016/0161-5890(83)90029-9.

Horák J. Regulations of sugar transporters: insights from yeast. Curr Genet. 2013:**59**(1–2):1–31. https://doi.org/10.1007/s00294-013-0388-8.

Hughes AL. The evolution of functionally novel proteins after gene duplication. Proc R Soc Lond B Biol Sci. 1994:**256**(1346): 119–124. https://doi.org/10.1098/rspb.1994.0058.

Hutzler M, Michel M, Kunz O, Kuusisto T, Magalhães F, Krogerus K, Gibson B. Unique brewing-relevant properties of a strain of *Saccharomyces jurei* isolated from ash (*Fraxinus excelsior*). Front Microbiol. 2021:**12**:645271. https://doi.org/10.3389/fmicb.2021.645271.

Jabłońska J, Tawfik DS. Innovation and tinkering in the evolution of oxidases. Protein Sci. 2022:**31**(5):e4310. https://doi.org/10.1002/pro.4310.

Janin J. Surface and inside volumes in globular proteins. Nature. 1979:**277**(5696):491–492. https://doi.org/10.1038/277491a0.

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021: **596**(7873):583–589. https://doi.org/10.1038/s41586-021-03819-2.

Kaback HR, Guan L. It takes two to tango: the dance of the permease. J Gen Physiol. 2019:**151**(7):878–886. https://doi.org/10.1085/jgp.201912377.

Kaltenbach M, Burke JR, Dindo M, Pabis A, Munsberg FS, Rabin A, Kamerlin SCL, Noel JP, Tawfik DS. Evolution of chalcone isomerase from a noncatalytic ancestor. Nat Chem Biol. 2018:**14**(6): 548–555. https://doi.org/10.1038/s41589-018-0042-3.

Kalyaanamoorthy S, Minh BQ, Wong TKF, Von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017:**14**(6):587–589. https://doi.org/10.1038/nmeth.4285.

Kasahara M, Shimoda E, Maeda M. Amino acid residues responsible for galactose recognition in yeast Gal2 transporter. J Biol Chem. 1997:**272**(27):16721–16724. https://doi.org/10.1074/jbc.272.27.16721.

Kasahara T, Ishiguro M, Kasahara M. Eight amino acid residues in transmembrane segments of yeast glucose transporter Hxt2 are required for high affinity transport. J Biol Chem. 2006:**281**(27):18532–18538. https://doi.org/10.1074/jbc.M602123200.

Kasahara T, Kasahara M. Tryptophan 388 in putative transmembrane segment 10 of the rat glucose transporter Glut1 is essential for glucose transport. J Biol Chem. 1998:**273**(44):29113–29117. https://doi.org/10.1074/jbc.273.44.29113.

Kasahara T, Kasahara M. Three aromatic amino acid residues critical for galactose transport in yeast Gal2 transporter. J Biol Chem. 2000:**275**(6):4422–4428. https://doi.org/10.1074/jbc.275.6.4422.

Kasahara T, Kasahara M. Identification of a key residue determining substrate affinity in the yeast glucose transporter Hxt7. J Biol Chem. 2010:**285**(34):26263–26268. https://doi.org/10.1074/jbc.M110.149716.

Kasahara T, Maeda M, Boles E, Kasahara M. Identification of a key residue determining substrate affinity in the human glucose transporter GLUT1. Biochim Biophys Acta. 2009:**1788**(5): 1051–1055. https://doi.org/10.1016/j.bbamem.2009.01.014.

Kasahara T, Maeda M, Ishiguro M, Kasahara M. Identification by comprehensive chimeric analysis of a key residue responsible for high affinity glucose transport by yeast HXT2. J Biol Chem. 2007:**282**(18):13146–13150. https://doi.org/10.1074/jbc.C700041200.

Kasahara T, Shimogawara K, Kasahara M. Crucial effects of amino acid side chain length in transmembrane segment 5 on substrate affinity in yeast glucose transporter Hxt7. Biochemistry. 2011:**50**(40):8674–8681. https://doi.org/10.1021/bi200958s.

Katoh K, Kuma KI, Toh H, Miyata T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res. 2005:**33**(2):511–518. https://doi.org/10.1093/nar/gki198.

Katoh K, Misawa K, Kuma KI, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002:**30**(14):3059–3066. https://doi.org/10.1093/nar/gkf436.

Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013:**30**(4):772–780. https://doi.org/10.1093/molbev/mst010.

Kauffman S, Levin S. Towards a general theory of adaptive walks on rugged landscapes. J Theor Biol. 1987:**128**(1):11–45. https://doi.org/10.1016/S0022-5193(87)80029-2.

Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015:**10**(6):845–858. https://doi.org/10.1038/nprot.2015.053

Khersonsky O, Tawfik DS. Enzyme promiscuity: a mechanistic and evolutionary perspective. Annu Rev Biochem. 2010:**79**(1):471–505. https://doi.org/10.1146/annurev-biochem-030409-143718.

Kim H, Oh EJ, Lane ST, Lee WH, Cate JHD, Jin YS. Enhanced cellobiose fermentation by engineered Saccharomyces cerevisiae expressing a mutant cellodextrin facilitator and cellobiose phosphorylase. J Biotechnol. 2018:**275**:53–59. https://doi.org/10.1016/j.jbiotec.2018.04.008.

Kriventseva EV, Kuznetsov D, Tegenfeldt F, Manni M, Dias R, Simão FA, Zdobnov EM. OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. Nucleic Acids Res. 2019:**47**(D1):D807–D811. https://doi.org/10.1093/nar/gky1053.

Kuang MC, Kominek J, Alexander WG, Cheng J-F, Wrobel RL, Hittinger CT. Repeated cis-regulatory tuning of a metabolic bottleneck gene during evolution. Mol Biol Evol. 2018:**35**(8): 1968–1981. https://doi.org/10.1093/molbev/msy102.

Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol. 1982:**157**(1):105–132. https://doi.org/10.1016/0022-2836(82)90515-0.

Langdon QK, Peris D, Eizaguirre JI, Opulente DA, Buh K V, Sylvester K, Jarzyna M, Rodríguez ME, Lopes CA, Libkind D, et al. Postglacial migration shaped the genomic diversity and global distribution of the wild ancestor of lager-brewing hybrids. PLoS Genet. 2020:**16**(4): e1008680. https://doi.org/10.1371/journal.pgen.1008680.

Lee SB, Tremaine M, Place M, Liu L, Pier A, Krause DJ, Xie D, Zhang Y, Landick R, Gasch AP, et al. Crabtree/Warburg-like aerobic xylose fermentation by engineered Saccharomyces cerevisiae. Metab Eng. 2021:**68**:119–130. https://doi.org/10.1016/j.ymben.2021.09.008.

Leffler EM, Band G, Busby GBJ, Kivinen K, Le QS, Clarke GM, Bojang KA, Conway DJ, Jallow M, Sisay-Joof F, et al. Resistance to malaria through structural variation of red blood cell invasion receptors. Science. 2017:**356**(6343):1140–1152. https://doi.org/10.1126/science.aam6393.

Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021:**49**(W1):W293–W296. https://doi.org/10.1093/nar/gkab301.

Li XC, Fay JC. Multiple changes underlie allelic divergence of CUP2 between Saccharomyces species. G3 (Bethesda). 2019:**9**(11): 3595–3600. https://doi.org/10.1534/g3.119.400616.

Libkind D, Hittinger CT, Valério E, Gonçalves C, Dover J, Johnston M, Gonçalves P, Sampaio JP. Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. Proc Natl Acad Sci U S A. 2011:**108**(35):14539–14544. https://doi.org/10.1073/pnas.1105430108.

Liu H, Maclean CJ, Zhang J. Evolution of the yeast recombination landscape. Mol Biol Evol. 2019:**36**(2):412–422. https://doi.org/10.1093/molbev/msy233.

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): 865–875. https://doi.org/10.1038/nrg1204.

Long M, Langley CH. Natural selection and the origin of jingwei, a chimeric processed functional gene in Drosophila. Science. 1993:**260**(5104):91–95. https://doi.org/10.1126/science.7682012.

Lunzer M, Miller SP, Felsheim R, Dean AM. Evolution: the biochemical architecture of an ancient adaptive landscape. Science. 2005:**310**(5747):499–501. https://doi.org/10.1126/science.1115649.

Lynch M. Simple evolutionary pathways to complex proteins. Protein Sci. 2005:**14**(9):2217–2225. https://doi.org/10.1110/ps.041171805.

Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. Mol Biol Evol. 2021:**38**(10): 4647–4654. https://doi.org/10.1093/molbev/msab199.

Marger MD, Saier MH. A major superfamily of transmembrane facilitators that catalyse uniport, symport and antiport. Trends Biochem Sci. 1993:**18**(1):13–20. https://doi.org/10.1016/0968-0004(93)90081-W.

Mefford HC, Trask BJ. The complex structure and dynamic evolution of human subtelomeres. Nat Rev Genet. 2002:**3**(2):91–102. https://doi.org/10.1038/nrg727.

Meger AT, Spence MA, Sandhu M, Matthews D, Chen J, Jackson CJ, Raman S. Rugged fitness landscapes minimize promiscuity in the evolution of transcriptional repressors. Cell Syst. 2024:**15**(4): 374–387.e6. https://doi.org/10.1016/j.cels.2024.03.002.

Méheust R, Bhattacharya D, Pathmanathan JS, McInerney JO, Lopez P, Bapteste E. Formation of chimeric genes with essential functions at the origin of eukaryotes. BMC Biol. 2018:**16**(1):30. https://doi.org/10.1186/s12915-018-0500-0.

Metzger BP, Park Y, Starr TN, Thornton JW. Epistasis facilitates functional evolution in an ancient transcription factor. Elife. 2024:**12**: RP88737. https://doi.org/10.7554/eLife.88737.3.

Meussdorfer F, Zarnkow M. Starchy raw materials. In: Esslinger HM, editor. Handbook of brewing: process, technology, markets. Weinheim: Wiley-VCH; 2009. p. 43–83.

Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R, Teeling E. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. Mol Biol Evol. 2020:**37**(5):1530–1534. https://doi.org/10.1093/molbev/msaa015.

Miyazaki K, Arnold FH. Exploring nonnatural evolutionary pathways by saturation mutagenesis: rapid improvement of protein function. J Mol Evol. 1999:**49**(6):716–720. https://doi.org/10.1007/PL00006593.

Nielsen R, Bustamante C, Clark AG, Glanowski S, Sackton TB, Hubisz MJ, Fledel-Alon A, Tanenbaum DM, Civello D, White TJ, et al. A scan for positively selected genes in the genomes of humans and chimpanzees. PLoS Biol. 2005:**3**(6):e170. https://doi.org/10.1371/journal.pbio.0030170.

Nijland JG, Driessen AJM. Engineering of pentose transport in Saccharomyces cerevisiae for biotechnological applications. Front Bioeng Biotechnol. 2020:**7**:464. https://doi.org/10.3389/fbioe.2019.00464.

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:**124**(2):503–510. https://doi.org/10.1111/jam.13670.

Nishikawa KK, Hoppe N, Smith R, Bingman C, Raman S. Epistasis shapes the fitness landscape of an allosteric specificity switch. Nat Commun. 2021:**12**(1):5562. https://doi.org/10.1038/s41467-021-25826-7.

Nurminsky DI, Nurminskaya M V, De Aguiar D, Hartl DL. Selective sweep of a newly evolved sperm-specific gene in *Drosophila*. Nature. 1998:**396**(6711):572–575. https://doi.org/10.1038/25126.

O'Donnell S, Yue J-X, Saada OA, Agier N, Caradec C, Cokelaer T, De Chiara M, Delmas S, Dutreux F, Fournier T, et al. Telomere-to-telomere assemblies of 142 strains characterize the genome structural landscape in *Saccharomyces cerevisiae*. Nat Genet. 2023:**55**(8): 1390–1399. https://doi.org/10.1038/s41588-023-01459-y.

Oh EJ, Jin YS. Engineering of *Saccharomyces cerevisiae* for efficient fermentation of cellulose. FEMS Yeast Res. 2020:**20**(1):foz089. https://doi.org/10.1093/femsyr/foz089.

Oh EJ, Kwak S, Kim H, Jin Y-S. Transporter engineering for cellobiose fermentation under lower pH conditions by engineered *Saccharomyces cerevisiae*. Bioresour Technol. 2017:**245**: 1469–1475. https://doi.org/10.1016/j.biortech.2017.05.138.

Opulente DA, LaBella AL, Harrison MC, Wolters JF, Liu C, Li Y, Kominek J, Steenwyk JL, Stoneman HR, VanDenAvond J, et al. Genomic factors shape carbon and nitrogen metabolic niche breadth across *Saccharomycotina* yeasts. Science. 2024:**384**: eadj4503. https://doi.org/10.1126/science.adj4503.

Ortlund EA, Bridgham JT, Redinbo MR, Thornton JW. Crystal structure of an ancient protein: evolution by conformational epistasis. Science. 2007:**317**(5844):1544–1548. https://doi.org/10.1126/science.1142819.

Pao SS, Paulsen IT, Saier MH. Major facilitator superfamily. Microbiol Mol Biol Rev. 1998:**62**(1):1–34. https://doi.org/10.1128/MMBR.62.1.1-34.1998.

Park Y, Metzger BPH, Thornton JW. Epistatic drift causes gradual decay of predictability in protein evolution. Science. 2022:**376**(6595): 823–830. https://doi.org/10.1126/science.abn6895.

Patthy L. Modular assembly of genes and the evolution of new functions. Genetica. 2003:**118**(2/3):217–231. https://doi.org/10.1023/A:1024182432483.

Paulsen PA, Custódio TF, Pedersen BP. Crystal structure of the plant symporter STP10 illuminates sugar uptake mechanism in monosaccharide transporter superfamily. Nat Commun. 2019:**10**(1): 407. https://doi.org/10.1038/s41467-018-08176-9.

Peris D, Ubbelohde EJ, Kuang MC, Kominek J, Langdon QK, Adams M, Koshalek JA, Hulfachor AB, Opulente DA, Hall DJ, et al. Macroevolutionary diversity of traits and genomes in the model yeast genus *Saccharomyces*. Nat Commun. 2023:**14**(1):690. https://doi.org/10.1038/s41467-023-36139-2.

Peter J, De Chiara M, Friedrich A, Yue J-X, Pflieger D, Bergström A, Sigwalt A, Barre B, Freel K, Llored A, et al. Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates. Nature. 2018: **556**(7701):339–344. https://doi.org/10.1038/s41586-018-0030-5.

Podgornaia AI, Laub MT. Pervasive degeneracy and epistasis in a protein-protein interface. Science. 2015:**347**(6222):673–677. https://doi.org/10.1126/science.1257360.

Poelwijk FJ, Kiviet DJ, Weinreich DM, Tans SJ. Empirical fitness landscapes reveal accessible evolutionary paths. Nature. 2007:**445**(7126): 383–386. https://doi.org/10.1038/nature05451.

Pokusaeva VO, Usmanova DR, Putintseva E V, Espinar L, Sarkisyan KS, Mishin AS, Bogatyreva NS, Ivankov DN, Akopyan A V, Avvakumov SY, et al. An experimental assay of the interactions of amino acids from orthologous sequences shaping a complex fitness landscape. PLoS Genet. 2019:**15**(4):e1008079. https://doi.org/10.1371/journal.pgen.1008079.

Pougach K, Voet A, Kondrashov FA, Voordeckers K, Christiaens JF, Baying B, Benes V, Sakai R, Aerts J, Zhu B, et al. Duplication of a promiscuous transcription factor drives the emergence of a new regulatory network. Nat Commun. 2014:**5**(1):4868. https://doi.org/10.1038/ncomms5868.

Projecto-Garcia J, Natarajan C, Moriyama H, Weber RE, Fago A, Cheviron ZA, Dudley R, McGuire JA, Witt CC, Storz JF. Repeated elevational transitions in hemoglobin function during the evolution of Andean hummingbirds. Proc Natl Acad Sci U S A. 2013:**110**(51): 20669–20674. https://doi.org/10.1073/pnas.1315456110.

Quistgaard EM, Löw C, Guettou F, Nordlund P. Understanding transport by the major facilitator superfamily (MFS): structures pave the way. Nat Rev Mol Cell Biol. 2016:**17**(2):123–132. https://doi.org/10.1038/nrm.2015.25.

Roach MJ, Borneman AR. New genome assemblies reveal patterns of domestication and adaptation across *Brettanomyces* (*Dekkera*) species. BMC Genomics. 2020:**21**(1):194. https://doi.org/10.1186/s12864-020-6595-z.

Rogers RL, Bedford T, Lyons AM, Hartl DL. Adaptive impact of the chimeric gene *Quetzalcoatl* in *Drosophila melanogaster*. Proc Natl Acad Sci U S A. 2010:**107**(24):10943–10948. https://doi.org/10.1073/pnas.1006503107.

Rogers RL, Hartl DL. Chimeric genes as a source of rapid evolution in *Drosophila melanogaster*. Mol Biol Evol. 2012:**29**(2):517–529. https://doi.org/10.1093/molbev/msr184.

Rose GD, Geselowitz AR, Lesser GJ, Lee RH, Zehfus MH. Hydrophobicity of amino acid residues in globular proteins. Science. 1985:**229**(4716): 834–838. https://doi.org/10.1126/science.4023714.

Rozpedowska E, Hellborg L, Ishchuk OP, Orhan F, Galafassi S, Merico A, Woolfit M, Compagno C, Piškur J. Parallel evolution of the make–accumulate–consume strategy in *Saccharomyces* and *Dekkera* yeasts. Nat Commun. 2011:**2**(1):302. https://doi.org/10.1038/ncomms1305.

Saier MH. Families of transmembrane sugar transport proteins. Mol Microbiol. 2000:**35**(4):699–710. https://doi.org/10.1046/j.1365-2958.2000.01759.x.

Saier MH, Reddy VS, Moreno-Hagelsieb G, Hendargo KJ, Zhang Y, Iddamsetty V, Lam KJK, Tian N, Russum S, Wang J, et al. The transporter classification database (TCDB): 2021 update. Nucleic Acids Res. 2021:**49**(D1):D461–D467. https://doi.org/10.1093/nar/gkaa1004.

Saier MH, Tran C V, Barabote RD. TCDB: the transporter classification database for membrane transport protein analyses and information. Nucleic Acids Res. 2006:**34**(90001):D181–D186. https://doi.org/10.1093/nar/gkj001.

Salema-Oom M, Valadão Pinto V, Gonçalves P, Spencer-Martins I. Maltotriose utilization by industrial *Saccharomyces* strains: characterization of a new member of the α-glucoside transporter family. Appl Environ Microbiol. 2005:**71**(9):5044–5049. https://doi.org/10.1128/AEM.71.9.5044-5049.2005.

Sarkisyan KS, Bolotin DA, Meer M V, Usmanova DR, Mishin AS, Sharonov G V, Ivankov DN, Bozhanova NG, Baranov MS, Soylemez O, et al. Local fitness landscape of the green fluorescent protein. Nature. 2016:**533**(7603):397–401. https://doi.org/10.1038/nature17995.

Shah P, McCandlish DM, Plotkin JB. Contingency and entrenchment in protein evolution under purifying selection. Proc Natl Acad Sci U S A. 2015:**112**(25):E3226–E3235. https://doi.org/10.1073/pnas.1412933112.

Shen X-X, Opulente DA, Kominek J, Zhou X, Steenwyk JL, Buh K V, Haase MAB, Wisecaver JH, Wang M, Doering DT, et al. Tempo and mode of genome evolution in the budding yeast subphylum. Cell. 2018:**175**(6):1533–1545.e20. https://doi.org/10.1016/j.cell.2018.10.023.

Smith JM. Natural selection and the concept of a protein space. Nature. 1970:**225**(5232):563–564. https://doi.org/10.1038/225563a0.

Smithers B, Oates M, Gough J. 'Why genes in pieces?'—revisited. Nucleic Acids Res. 2019:**47**(10):4970–4973. https://doi.org/10.1093/nar/gkz284.

Soylemez O, Kondrashov FA. Estimating the rate of irreversibility in protein evolution. Genome Biol Evol. 2012:**4**(12):1213–1222. https://doi.org/10.1093/gbe/evs096.

Sprouffske K, Wagner A. Growthcurver: an R package for obtaining interpretable metrics from microbial growth curves. BMC

Bioinformatics. 2016:17(1):172. https://doi.org/10.1186/s12859-016-1016-7.

Srikant S, Gaudet R, Murray AW. Selecting for altered substrate specificity reveals the evolutionary flexibility of ATP-binding cassette transporters. Curr Biol. 2020:30(9):1689–1702.e6. https://doi.org/10.1016/j.cub.2020.02.077.

Stambuk BU, Araujo PS. Kinetics of active alpha-glucoside transport in Saccharomyces cerevisiae. FEMS Yeast Res. 2001:1(1):73–78. https://doi.org/10.1111/j.1567-1364.2001.tb00015.x.

Stambuk BU, Batista AS, De Araujo PS. Kinetics of active sucrose transport in Saccharomyces cerevisiae. J Biosci Bioeng. 2000:89(2):212–214. https://doi.org/10.1016/S1389-1723(00)88742-3.

Stambuk BU, Silva MA, Panek AD, Araujo PS. Active alpha-glucoside transport in Saccharomyces cerevisiae. FEMS Microbiol Lett. 1999:170(1):105–110. https://doi.org/10.1111/j.1574-6968.1999.tb13361.x.

Stanke M, Diekhans M, Baertsch R, Haussler D. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. Bioinformatics. 2008:24(5):637–644. https://doi.org/10.1093/bioinformatics/btn013.

Starr TN, Flynn JM, Mishra P, Bolon DNA, Thornton JW. Pervasive contingency and entrenchment in a billion years of Hsp90 evolution. Proc Natl Acad Sci U S A. 2018:115(17):4453–4458. https://doi.org/10.1073/pnas.1718133115.

Starr TN, Picton LK, Thornton JW. Alternative evolutionary histories in the sequence space of an ancient protein. Nature. 2017:549(7672):409–413. https://doi.org/10.1038/nature23902.

Starr TN, Thornton JW. Epistasis in protein evolution. Protein Sci. 2016:25(7):1204–1218. https://doi.org/10.1002/pro.2897.

Steindel PA, Chen EH, Wirth JD, Theobald DL. Gradual neofunctionalization in the convergent evolution of trichomonad lactate and malate dehydrogenases. Protein Sci. 2016:25(7):1319–1331. https://doi.org/10.1002/pro.2904.

Sun L, Zeng X, Yan C, Sun X, Gong X, Rao Y, Yan N. Crystal structure of a bacterial homologue of glucose transporters GLUT1–4. Nature. 2012:490(7420):361–366. https://doi.org/10.1038/nature11524.

Townsend JP, Nielsen KM, Fishert DS, Hartl DL. Horizontal acquisition of divergent chromosomal DNA in bacteria: effects of mutator phenotypes. Genetics. 2003:164(1):13–21. https://doi.org/10.1093/genetics/164.1.13.

Trichez D, Knychala MM, Figueiredo CM, Alves SL, da Silva MA, Miletti LC, de Araujo PS, Stambuk BU. Key amino acid residues of the AGT1 permease required for maltotriose consumption and fermentation by Saccharomyces cerevisiae. J Appl Microbiol. 2019:126(2):580–594. https://doi.org/10.1111/jam.14161.

Turner TL, Kim H, Kong II, Liu J-J, Zhang G-C, Jin Y-S. Engineering and evolution of Saccharomyces cerevisiae to produce biofuels and chemicals. In: Advances in biochemical engineering/biotechnology. Vol. 162. Cham, Switzerland: Springer; 2016. p. 175–215.

Voordeckers K, Brown CA, Vanneste K, van der Zande E, Voet A, Maere S, Verstrepen KJ. Reconstruction of ancestral metabolic enzymes reveals molecular mechanisms underlying evolutionary innovation through gene duplication. PLoS Biol. 2012:10(12):e1001446. https://doi.org/10.1371/journal.pbio.1001446.

Wang SC, Davejan P, Hendargo KJ, Javadi-Razaz I, Chou A, Yee DC, Ghazi F, Lam KJK, Conn AM, Madrigal A, et al. Expansion of the major facilitator superfamily (MFS) to include novel transporters as well as transmembrane-acting enzymes. Biochim Biophys Acta. 2020:1862(9):183277. https://doi.org/10.1016/j.bbamem.2020.183277.

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):1294–1301. https://doi.org/10.1093/oxfordjournals.molbev.a026413.

Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, et al. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018a:46(W1):W296–W303. https://doi.org/10.1093/nar/gky427.

Waterhouse RM, Seppey M, Simao FA, Manni M, Ioannidis P, Klioutchnikov G, Kriventseva E V, Zdobnov EM. BUSCO applications from quality assessments to gene prediction and phylogenomics. Mol Biol Evol. 2018b:35(3):543–548. https://doi.org/10.1093/molbev/msx319.

Weinreich DM, Delaney NF, DePristo MA, Hartl DL. Darwinian evolution can follow only very few mutational paths to fitter proteins. Science. 2006:312(5770):111–114. https://doi.org/10.1126/science.1123539.

Weinreich DM, Lan Y, Wylie CS, Heckendorn RB. Should evolutionary geneticists worry about higher-order epistasis? Curr Opin Genet Dev. 2013:23(6):700–707. https://doi.org/10.1016/j.gde.2013.10.007.

Weinreich DM, Watson RA, Chao L. Perspective: sign epistasis and genetic costraint on evolutionary trajectories. Evolution. 2005:59:1165–1174. https://doi.org/10.1111/j.0014-3820.2005.tb01768.x.

Wheeler LC, Harms MJ. Were ancestral proteins less specific? Mol Biol Evol. 2021:38(6):2227–2239. https://doi.org/10.1093/molbev/msab019.

Wheeler LC, Lim SA, Marqusee S, Harms MJ. The thermostability and specificity of ancient proteins. Curr Opin Struct Biol. 2016:38:37–43. https://doi.org/10.1016/j.sbi.2016.05.015.

Wieczorke R, Krampe S, Weierstall T, Freidel K, Hollenberg CP, Boles E. Concurrent knock-out of at least 20 transporter genes is required to block uptake of hexoses in Saccharomyces cerevisiae. FEBS Lett. 1999:464(3):123–128. https://doi.org/10.1016/S0014-5793(99)01698-1.

Wong ED, Miyasato SR, Aleksander S, Karra K, Nash RS, Skrzypek MS, Weng S, Engel SR, Cherry JM. Saccharomyces genome database update: server architecture, pan-genome nomenclature, and external resources. Genetics. 2023:224(1):191. https://doi.org/10.1093/genetics/iyac191.

Woolfit M, Rozpędowska E, Piškur J, Wolfe KH. Genome survey sequencing of the wine spoilage yeast Dekkera (Brettanomyces) bruxellensis. Eukaryot Cell. 2007:6(4):721–733. https://doi.org/10.1128/EC.00338-06.

Wright S. 1932. The roles of mutation, inbreeding, crossbreeding, and selection in evolution. Proceedings of the 6th International Congress of Genetics. Vol. 1. Ithaca, NY. p. 356–366.

Wright S. Evolution in Mendelian populations. Genetics. 1931:16(2):97–159. https://doi.org/10.1093/genetics/16.2.97.

Wu NC, Dai L, Olson CA, Lloyd-Smith JO, Sun R. Adaptation in protein fitness landscapes is facilitated by indirect paths. Elife. 2016:5:e16965. https://doi.org/10.7554/eLife.16965.

Xia X, Li WH. What amino acid properties affect protein evolution? J Mol Evol. 1998:47(5):557–564. https://doi.org/10.1007/PL00006412.

Xie VC, Pu J, Metzger BPH, Thornton JW, Dickinson BC. Contingency and chance erase necessity in the experimental evolution of ancestral proteins. Elife. 2021:10:1–87. https://doi.org/10.7554/eLife.67336.

Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER suite: protein structure and function prediction. Nat Methods. 2015:12(1):7–8. https://doi.org/10.1038/nmeth.3213.

Yi X, Dean AM. Adaptive landscapes in the age of synthetic biology. Mol Biol Evol. 2019:36(5):890–907. https://doi.org/10.1093/molbev/msz004.

Yokoyama S, Tada T, Zhang H, Britt L. Elucidation of phenotypic adaptations: molecular analyses of dim-light vision proteins in vertebrates. Proc Natl Acad Sci U S A. 2008:105(36):13480–13485. https://doi.org/10.1073/pnas.0802426105.

Young EM, Tong A, Bui H, Spofford C, Alper HS. Rewiring yeast sugar transporter preference through modifying a conserved protein motif. Proc Natl Acad Sci U S A. 2014:111(1):131–136. https://doi.org/10.1073/pnas.1311970111.

Yue J-X, Li J, Aigrain L, Hallin J, Persson K, Oliver K, Bergström A, Coupland P, Warringer J, Lagomarsino MC, et al. Contrasting evolutionary genome dynamics between domesticated and

wild yeasts. Nat Genet. 2017:**49**(6):913–924. https://doi.org/10.1038/ng.3847.

Zastrow CR, Hollatz C, de Araujo PS, Stambuk BU. Maltotriose fermentation by *Saccharomyces cerevisiae*. J Ind Microbiol Biotechnol. 2001:**27**(1):34–38. https://doi.org/10.1038/sj.jim.7000158.

Zhang J, Dean AM, Brunet F, Long M. Evolving protein functional diversity in new genes of *Drosophila*. Proc Natl Acad Sci U S A. 2004:**101**(46):16246–16250. https://doi.org/10.1073/pnas.0407066101.

Zhang J, Zhang YP, Rosenberg HF. Adaptive evolution of a duplicated pancreatic ribonuclease gene in a leaf-eating monkey. Nat Genet. 2002:**30**(4):411–415. https://doi.org/10.1038/ng852.