Genomic signatures of innovation and selection in the extremotolerant yeast *Kluyveromyces marxianus*

Kaylee E. Christensen[1], Aniliese Deal[1,2], Abel Duarte[1], Zhenzhen Ma[1,3], Judith L. Edwards[1], Rachel B. Brem[1]*

[1]Department of Plant and Microbial Biology, University of California, Berkeley, Berkeley, CA, 94720

[2]Graduate Group in Biophysics, University of California, Berkeley, Berkeley, CA, 94720

[3]Current address: Department of Biology, Stanford University, Stanford, CA, 94305

*To whom correspondence should be addressed (rbrem@berkeley.edu)

Abstract

Extremophiles can be the product of millions of years of evolutionary engineering and refinement. The underlying mechanisms can be quite distinct from the ones operating at earlier stages of trait innovation. In this work, we have developed the compost yeast *Kluyveromyces marxianus,* which diverged from its closest relative >20 million years ago, as a model for interspecies comparative biology and genomics. We applied a battery of growth assays to species of the *Kluyveromyces* genus and found that *K. marxianus* outperformed its relatives in a battery of heat and chemical stress conditions. We then generated and analyzed genomes from across the genus, to find derived genetic features associated with, and potentially causal for, *K. marxianus* traits. We found robust expansions in gene families in the *K. marxianus* genome, most notably among genes annotated as transmembrane transporters and in metabolism. In molecular-evolution tests, we identified adaptive protein variants at hundreds of genes, among which plasma membrane transporters were over-represented. Together, these signals enable a model for the molecular mechanisms and evolutionary pressures underlying *K. marxianus* traits, including gains in transporter function mediating stress resistance, and metabolic variants contributing to its capacity for rapid growth in challenging conditions. Such oligogenic architectures may be the rule rather than the exception in phenotypes that have evolved over long timescales.

Introduction

A central goal of evolutionary biology is to understand trait diversity in the natural world. In many cases, a character of ecological or industrial interest defines a given eukaryotic species, distinguishing it from its relatives. Elucidating the mechanisms of any such phenotype across

species boundaries poses unique technical challenges. As a consequence, against a backdrop of landmark case studies mapping genotype to phenotype over deep trait divergences (Weiss and Brem 2019; Smith et al. 2020; Masly and Azom 2022), principles remain at a premium. What kinds of phenotypes tend to fix in species as they split from one another? How complex is the genetics, and what are the causal genes and variants? Addressing these questions requires detailed study of the physiology of ancient traits and their mechanisms.

Single-celled fungi can serve as a powerful model for evolutionary genetics. *Kluyveromyces,* a genus of budding yeasts, are well-suited to this paradigm given their ~25-million-year divergence (Shen et al. 2018) and wide-ranging ecology. *Kluyveromyces* have been isolated from a range of terrestrial and aquatic environments but are best characterized for roles in food and beverage fermentations (Kurtzman et al. 2011), including elegant evolutionary analyses of the origin of lactases in dairy-associated lineages (Varela, Puricelli, Ortiz-Merino, et al. 2019; Friedrich et al. 2023). An extensive observational literature has focused on the growth of one species in the genus, *K. marxianus,* at temperatures exceeding 45°C, motivated by applications in the bioproduction industry (Lane et al. 2011; Lertwattanasakul et al. 2013; Mejía-Barajas et al. 2017; Madeira-Jr and Gombert 2018; Fu et al. 2019; Hoshida et al. 2020; Wu et al. 2020; Kosaka et al. 2022; Flores-Cosío et al. 2024). The evolutionary driver for the latter trait is unknown, though *K. marxianus* has been isolated from self-heating agricultural and domestic compost (Suzuki et al. 2014; Avchar et al. 2022; Verma et al. 2023). Likewise, mechanisms underlying unique *K. marxianus* traits have remained incompletely understood (Nespolo et al. 2020); potential clues have emerged from comparison of this species to distantly related yeasts, including reports of resistance to chemical stress (Saini et al. 2017; Linder 2018; Illarionov et al. 2021) and expansions in sugar and potassium importer genes (Varela, Puricelli, Montini, et al. 2019; Papouskova et al. 2024).

64 We set out to use a comparative approach to trace in detail how and why *K. marxianus* diverged
65 from its relatives. We first catalogued stress resistance characters that distinguished *K.*
66 *marxianus* from the rest of the species of its clade. We then applied genome-content and
67 molecular-evolution analyses to find unique genomic features in *K. marxianus*. The results
68 provided insights into the mechanisms and evolutionary forces driving the *K. marxianus* trait
69 syndrome.
70

71 # Results

72

73 ***K. marxianus* exhibits a stress resistance syndrome unique in its genus**

74

75 With the ultimate goal of understanding evolutionary innovations in *Kluyveromyces*, we first
76 sought to survey stress resistance traits across the genus. We profiled the growth in liquid
77 culture of four species associated with terrestrial habitats (*K. marxianus, K. lactis, K.*
78 *wickerhamii*, and *K. dobzhanskii*) and three aquatic species (*K. aestuarii*, *K. siamensis*, and *K.*
79 *nonfermentans*) in a panel of media conditions (Figure 1 and Table S1). At 42°C, *K. marxianus*
80 exhibited a qualitative growth advantage over its relatives, to an extent far beyond the pattern in
81 a permissive control condition (rich medium at 28°C; Figure 1B); genetically distinct *K.*
82 *marxianus* isolates phenocopied each other with respect to high-temperature growth (Figure
83 S2), as expected (Lane et al. 2011). In ethanol, caffeine, the DNA intercalator propidium iodide,
84 and the DNA damage agent methyl methanesulfonate (MMS), *K. marxianus* likewise
85 outperformed the other species of the genus (Figure 1B). Our profiling also identified other
86 conditions in which *K. marxianus* did not stand out with respect to growth (Figure S1). We
87 conclude that *K. marxianus* is unique in the genus for its resistance to a broad subset of abiotic
88 challenges, including high temperature.

89

90 ***K. marxianus* resistance phenotypes are unique for their dependence on growth phase**
91 **but not oxygen**

92

93 For many microbes, when cells face nutrient exhaustion, they shut down exponential (log-
94 phase) growth and enter the quiescent, extra-durable state called stationary phase. The
95 difference between *K. marxianus* and its relatives in stress resistance that we had noted across
96 a growth timecourse (Figure 1) might also manifest in the stationary state or, alternatively, its
97 mechanism could be tied to active cell division. To distinguish between these possibilities, we
98 focused on thermotolerance as a test case. We used a viability assay in which we inoculated a
99 given liquid culture into rich medium; incubated for growth at the permissive temperature until
100 nutrient exhaustion; subjected this stationary-phase culture to a short heat shock; and used it to
101 seed colonies for growth at the permissive temperature on solid medium (Figure 2A). Results
102 revealed only modest differences between species in heat shock resistance in this stationary-
103 phase heat treatment paradigm, with no particular advantage for *K. marxianus* (Figure 2B). By
104 contrast, when we treated log-phase cultures with heat shock, *K. marxianus* exhibited much
105 higher viability than did the other species of the genus (Figure 2C), paralleling its advantage
106 during chronic high-temperature exposure (Figure 1B). No such differences were apparent in a
107 permissive 28°C control experiment (Figure S3). We conclude that for high temperature, the *K.*
108 *marxianus* advantage only manifests when cells are actively dividing, and that cell death, rather
109 than slow or arrested growth, mediates the defects of the rest of the genus.

110

111 To extend our analysis of cell death as the underlying mechanism of species differences in
112 stress, we next assayed viability in log-phase cultures across a panel of representative
113 stresses, focusing on the stress-sensitive species *K. lactis* as a point of comparison for *K.*
114 *marxianus* (Figure 1). The latter exhibited a survival advantage in each stress condition (Figure

115 S4). The difference manifested regardless of oxygen availability, though both species performed
116 better in normoxia (Figure S4), as expected (Lertwattanasakul et al. 2013; Mejía-Barajas et al.
117 2017; Madeira-Jr and Gombert 2018; Fu et al. 2019; Hoshida et al. 2020; Kosaka et al. 2022).
118 Together, these trends from assays of viability establish the resistance character of *K.*
119 *marxianus* as a lineage-unique gain of mechanisms that protect cells from lethal failures of the
120 proliferation program under stress.
121
122 **Transcriptional remodeling by *K. lactis*, but not *K. marxianus*, in stress**
123
124 To learn more about the state of *Kluyveromyces* cells under stress, we turned to an expression
125 profiling approach, for which we again chose *K. lactis* as a model stress-sensitive species for
126 comparison against *K. marxianus*. We surveyed the transcriptomes of the two species in
127 representative treatments where we had observed divergent growth phenotypes—high
128 temperature, caffeine, MMS, and ethanol—and in the permissive control of rich medium at 28°C
129 (Table S2). In principal component analyses, the strongest contributor to variation across the
130 data set was a program of housekeeping genes expressed more highly in *K. lactis* than *K.*
131 *marxianus* in all conditions, including controls (Figure 3A, Figure S5A, and Table S2). Also
132 salient on a global scale was a suite of genes that *K. lactis* turned up and down depending on
133 treatment; the *K. marxianus* transcriptome was much less volatile (Figure 3A, Figure S5A, and
134 Table S2).
135
136 For a more fine-grained analysis, we tested Gene Ontology groups for coherent, stress-evoked
137 expression change (Table S3). The results revealed housekeeping gene groups induced in
138 stress by *K. lactis*, including ribosomal RNA processing and protein folding (Figure 3B, Figure
139 S5B, and Table S3). Additionally, in a few cases, regulons in *K. lactis* and *K. marxianus*
140 paralleled each other, notably repression of glycolysis and induction of the electron transport

141 chain in some stresses (Figure 3C and Figure S5C). Taken together, however, our
142 transcriptional analyses established programs of elevated expression by *K. lactis*—in stress and
143 in untreated controls—as its major regulatory distinction from *K. marxianus*. Such a pattern
144 would be expected if *K. lactis* underwent more functional failures during growth in our panel of
145 conditions, setting off compensatory expression responses downstream.
146
147 **Unique transporter and metabolic gene duplications in the *K. marxianus* genome**
148
149 In continued pursuit of mechanisms underlying traits that distinguished *K. marxianus* from its
150 relatives, we reasoned that unique genomic features in this species could serve as a source for
151 potential hypotheses. As a first strategy to find genomic novelties in *K. marxianus*, we focused
152 on species-specific gene duplications in paralogous gene families. From annotated reference
153 genomes for six *Kluyveromyces* species (*K. marxianus, K. lactis, K. wickerhamii*, *K.*
154 *dobzhanskii*, *K. aestuarii*, and *K. nonfermentans*; Figure 1A), plus outgroups from nearby clades
155 in the Saccharomycotina subphylum, we identified gene families and the members detectable in
156 each species in turn with the CAFE 5 package (Mendes et al. 2021). This approach identified 16
157 gene families with inferred gain or loss events along the *K. marxianus* lineage at rates that were
158 unusual relative to the average across the species set (Figure 4A and Table S4), including
159 families subject to tandem duplications and those with members across the *K. marxianus*
160 chromosomes (Figure 4B and Table S4). Of these, 13 were families with extra copies in *K.*
161 *marxianus* that could be classified based on domain annotations as functioning in metabolism or
162 plasma membrane transport, the latter including both nutrient importers and xenobiotic efflux
163 pumps (Figure 4A). These data make clear that gene family expansions are prevalent in the *K.*
164 *marxianus* genome, providing a line of genomic evidence that highlights small-molecule
165 transport and metabolism as candidate mechanisms of *K. marxianus* phenotypes.
166

167 **Widespread evidence for positive selection on amino acid alleles in *K. marxianus***

168

169 As a further means to identify genomic correlates of *K. marxianus* phenotypes, we pursued
170 molecular-evolution analyses of amino acid variation on a per-gene basis. We reasoned that
171 some phenotypes unique to *K. marxianus* could have evolved as a product of adaptation to
172 abiotic stress in challenging ecological niches (including in compost), and that at relevant loci,
173 we could detect evidence for positive selection on protein-coding sequences. For a phylogenetic
174 test toward this end, at each gene in turn, we aligned orthologs from reference genomes for our
175 six focal *Kluyveromyces* species, to which we fit a model in the PAML package (Yang 2007)
176 specifying positive selection along the *K. marxianus* lineage for some codons, with all other
177 codons and branches evolving under selective constraint. This model was justified, over and
178 above a null model with no selection, for a remarkable 525 genes (Figure 5A, 5C, and 5D and
179 Table S5). By contrast, controls focused on the *K. lactis* branch identified just five genes with
180 signal for positive selection (Figure 5A and Table S5), supporting the inference of a history of
181 particularly avid adaptation at the amino acid level in *K. marxianus*.

182

183 For most gene hits from our PAML branch-site tests focused on *K. marxianus*, a small number
184 of codons per gene were classified as targets of positive selection (Figure 5B). We expected
185 that, if the derived alleles at each such site had undergone a selective sweep through *K.*
186 *marxianus* populations, we could detect it across strains of this species. To test this, we carried
187 out short-read resequencing of the genomes of five wild *K. marxianus* isolates (Table S1) and
188 used the resulting reads alongside our reference genome of the Km31 *K. marxianus* strain as
189 input to a pipeline for variant calling, pseudogenome construction, and sequence alignment. We
190 then inspected these alignments at each site where the *K. marxianus* allele had arisen under
191 selection according to our analysis of reference genomes with PAML. At 99% of the latter, we
192 detected complete conservation across *K. marxianus* strains (Figure 5C and Table S5). These

193 phylogenetic and resequencing data are consistent with the scenario of positive selection in *K.*
194 *marxianus* in hundreds of proteins.
195
196 **Enrichment for amino acid variation between *Kluyveromyces* species in transporter**
197 **genes**
198
199 To complement our phylogenetic analyses, we next applied population-based molecular-
200 evolution tests that used our *K. marxianus* resequencing data and a population-genomic
201 resource for *K. lactis* (Friedrich et al. 2023). We implemented McDonald-Kreitman molecular
202 evolution paradigms (McDonald and Kreitman 1991) to assess evidence for differences in
203 selective pressure between the species on a given protein sequence, as follows. For each
204 coding region we tabulated sites of non-synonymous and synonymous divergence between *K.*
205 *marxianus* and *K. lactis*, and likewise for polymorphic variants within the species (Table S6).
206 Using these counts in a McDonald-Kreitman test on a per-gene basis, we identified 11 genes
207 with a significant excess of interspecies amino acid divergence across the coding region
208 (relative to polymorphism and for synonymous variation; Figure 5D and Table S6). Thus,
209 relative to the branch-site phylogenetic method (Figure 5A), the McDonald-Kreitman scheme
210 was under-powered in this system, as expected if most loci only had a few amino acid targets of
211 adaptation between *K. marxianus* and *K. lactis* (Figure 5B). But among the gene hits that did
212 emerge from the McDonald-Kreitman test, each represented an inference of accelerated protein
213 divergence between the species that complemented our phylogenetic results; four such loci had
214 also been top-scoring in our phylogenetic framework, two annotated in metabolism and two in
215 transmembrane transport (Figure 5C and 5D).
216
217 Separately, we developed a genomic version of the McDonald-Kreitman paradigm, in which we
218 tabulated the neutrality index (NI) for a given gene as non-synonymous divergence normalized

219 for polymorphism and for synonymous variation, and we evaluated significance of this metric
220 across cohorts of genes of similar function via resampling (Figure 5E and Table S7). For
221 metabolic gene groups with extreme NIs in this test, focused analyses indicated that the signal
222 was driven by low polymorphism, reflecting a history of purifying selection (Table S7). By
223 contrast, for a group of transmembrane transporters, the individual components of the NI metric
224 followed the pattern indicative of *bona fide* positive selection: non-synonymous variation
225 between species was high in the genes of the group and polymorphism levels were on par with
226 the genomic average (Figure 5E and Table S7). This selection signal in transporter genes as a
227 group dovetailed with our observation of individual transporters emerging as top-scoring in
228 single-gene McDonald-Kreitman tests as phylogenetic analyses (Figure 5D); it also echoed our
229 observation of transporter gene family expansions in *K. marxianus* (Figure 4). As such, our
230 complete set of genomic analyses attest to a complex history of selection on loci throughout the
231 *K. marxianus* genome, and they uncover innovations in transporter function as a particularly
232 compelling candidate mechanism of unique traits in this species.
233

234 ## Discussion

235

236 In this work, we have established *Kluyveromyces* as a genomic and genetic model for the study
237 of species-unique traits. Our results have revealed attributes that distinguish *K. marxianus* from
238 the rest of its genus: stress resistance in growth assays, expression stability in the face of
239 environmental change, and sequence evidence for gene-content and amino-acid variation.
240 These observations shed new light on the biology and evolution of *K. marxianus*,
241 complementing classic trait compendia (Kurtzman et al. 2011) and the modern literature (Lane
242 and Morrissey 2010; Lane et al. 2011; Madeira-Jr and Gombert 2018). In terms of genetics, the
243 alleles we have found to be unique in *K. marxianus* represent candidates for the components

that nature would have used to build the trait syndrome in this species and that would be sufficient to reconstitute it in an exogenous background—as opposed to the handful of previous studies that have reported genes necessary for thermotolerance in *K. marxianus* (Falcão Moreira et al. 1998; Lertwattanasakul et al. 2013; Wu et al. 2020; Montini et al. 2022).

Most pervasive in our genomic sequence scans was the trend of expanded transmembrane transporter gene families, and new amino acid alleles in single-copy transporter genes, in the *K. marxianus* genome relative to the rest of its clade. These include gains by *K. marxianus* in sugar and potassium transporter repertoires, which were previously analyzed in comparisons with distant fungal relatives (Varela, Puricelli, Montini, et al. 2019; Papouskova et al. 2024) and which we can now date to the divergence of *K. marxianus* from its sister species. Our results also uncovered alleles unique to *K. marxianus* in many other ion and nutrient importers that function at the plasma membrane, as well as xenobiotic exporters and the transporter regulatory machinery called the eisosome. Such changes could contribute to the stress resistance phenotypes we have seen in *K. marxianus* if they improve the selectivity of nutrient importers or the avidity of toxin exporters. Alternatively, transporter changes in *K. marxianus* could represent compensation for a separate, as yet unknown heightened barrier function at the plasma membrane and/or cell wall: any such extra protection could boost stress resistance at the expense of passive nutrient acquisition (Ferenci 2016) and drive compensatory changes to improve the machinery for active transport.

That said, we observed signals of evolutionary innovation by *K. marxianus* in genes of many functions beyond transmembrane transport. We thus picture a multifaceted mechanism for resistance traits in this system. Under one extreme model, the complement of new gene copies and new amino acid alleles that we have found in *K. marxianus* could work together to mediate a largely unified, global stress-resistance program. Perhaps more plausibly, subsets of these

270 variants could have more specific phenotypic roles, if *K. marxianus* evolved different
271 mechanisms for resistance to different insults. Indeed, hit loci from our analyses of the *K.*
272 *marxianus* genome span a much wider range of annotations than the variant cell growth genes
273 mapped in previous studies of *S. cerevisiae* thermotolerance (Weiss et al. 2018; AlZaben et al.
274 2021; Abrams et al. 2022)—likely reflecting the evolution of resistance to a bigger swath of
275 stresses by *K. marxianus.*
276
277 If the trait syndrome in *K. marxianus* were mediated by hundreds of unlinked loci, it would
278 conform to the view of genetic complexity as the rule rather than the exception for trait variation,
279 across the tree of life (Rockman 2012; Boyle et al. 2017). Even the novel gene copies and
280 amino acid alleles we have tabulated in the *K. marxianus* genome could represent an
281 underestimate of the full genetic architecture of traits in this species, given the potential of false
282 negatives in our approach. In other systems, signal for selection on such a scale has been
283 reported in the study of major life transitions (*e.g.* the evolution of marine mammals from
284 terrestrial ancestors (Chikina et al. 2016) and snakes from tetrapods (Castoe et al. 2013)), and
285 in analyses of convergent traits in species complexes (Smith et al. 2020), from tomato (Pease et
286 al. 2016) to long-lived fishes (Kolora et al. 2021) and eusocial bees (Woodard et al. 2011). The
287 ecological pressures on the *K. marxianus* ancestor remain poorly established. Though well-
288 recognized as an industrial dairy yeast (Prado et al. 2015; Gethins et al. 2016; Ortiz-Merino et
289 al. 2018; Varela, Puricelli, Ortiz-Merino, et al. 2019; Yazdi et al. 2022), *K. marxianus* was likely
290 domesticated for this purpose fairly recently, as inferred for *K. lactis* (Friedrich et al. 2023). And
291 though *K. marxianus* can also be isolated from rotting fruit (Williams et al. 2019; Al-Qaysi et al.
292 2021), it may not have specialized narrowly to such substrates: a neighboring clade that
293 includes *Saccharomyces cerevisiae* produces more ethanol than do any of the *Kluyveromyces*
294 species (Nespolo et al. 2020), suggesting a weaker drive in the latter toward this putative
295 mechanism to kill off microbial competitors in high-sugar environments. Among microbial-

296 ecology findings, the strongest clue to the origin of unique phenotypes in *K. marxianus* lies in its
297 prevalence in decomposing plant material other than fruit, including leaf litter and pine needles
298 (Spurley et al. 2022), cow dung (Avchar et al. 2022), industrially processed cocoa beans
299 (Baptista et al. 2021), agave (Chacón-Vargas et al. 2020), and sugar cane (Anderson et al.
300 1986; Cernak et al. 2018) as well as industrial and domestic compost (Suzuki et al. 2014;
301 Avchar et al. 2022; Verma et al. 2023). Plant defense compounds in decomposing material are
302 a compelling candidate driver for the rise of resistance traits in *K. marxianus* (Baptista and
303 Domingues 2022) since they can persist for days in a leaf or stem after abscission (Burton
304 1982; Goodspeed et al. 2013; Shimada et al. 2015). Furthermore, some evidence for low pH
305 deep in compost piles (Rainey and Covey 2016) suggests that this condition could also have
306 contributed to *K. marxianus* evolution. In this scenario, the phenotypic syndrome of *K.*
307 *marxianus* would involve not only resistance to the stresses of the decaying plant environment,
308 but also the rapid utilization of simple sugars liberated early in plant breakdown, mediated by
309 the new copies and new alleles of metabolic genes we have noted in the *K. marxianus* genome.
310 As *K. marxianus* lacks the capacity to degrade lignocellulosic plant cell walls, these would be
311 left to other fungal saprotrophs acting later in decomposition.
312
313 We thus envision an ecological model in which *K. marxianus* originally evolved as a plant decay
314 specialist, with extant strains coming to occupy a broad range of niches, including soil and
315 human and animal material (Rij 2013; Ortiz-Merino et al. 2018; Zhang et al. 2021). Given that
316 we and others have seen conservation of resistance traits across *K. marxianus* strains (Lane et
317 al. 2011; Madeira-Jr and Gombert 2018; Lappe-Oliveras et al. 2023), we infer that unique
318 capacities arose early in *K. marxianus,* sometime after *K. marxianus* originated as a species,
319 ~25 million years ago (Shen et al. 2018), and persisted during its radiation into modern
320 populations.
321

322 Assuming that hundreds of species-unique variants across the genome prove to explain the *K.*
323 *marxianus* trait syndrome, they could have accumulated over a very long time. Current theory
324 (Jain and Stephan 2017; Hayward and Sella 2022) would posit that, along the *K. marxianus*
325 lineage, a first advantage arose in ecologically relevant stress conditions as a product of a few
326 major-effect loci, followed later by the accumulation of many more variants with smaller adaptive
327 effects. Dynamics of this kind, assuming they emerge as generally relevant, could unify the
328 current understanding of genetically simple, recent, intraspecies (local) adaptations (Bell 2009)
329 with that of more ancient traits. Empirical validation of this model, in lower and higher
330 organismal systems, represents an ongoing goal of interspecific genetics.
331

332 # Methods

333

334 **1. Strains**

335

336 Strains used in phenotyping experiments and population-genomic analyses are listed in Table
337 S1.

338

339 **2. Liquid growth phenotyping of wild-type *Kluyveromyces***

340

341 Liquid culture experiments in stress and standard-condition controls in Figure 1 used *K.*
342 *marxianus* Km31 (Cernak et al. 2018); *K. lactis* yLB72 (Sorrells et al. 2015); and the strains of
343 *K. dobzhanskii, K. wickerhamii, K. nonfermentans, K. aestuarii*, and *K. siamensis* listed in Table
344 S1. For a given species and condition, we streaked a frozen stock onto yeast peptone dextrose
345 (YPD) agar plates, incubated for two days at 28°C, picked a single colony, inoculated into 5 mL
346 liquid YPD, and incubated with shaking at 28°C for 16-18 hours. We call this the first pre-

347 culture. We diluted, into fresh liquid YPD, an inoculum of the first pre-culture at the volume
348 required to attain an optical density (OD_{600}) of 0.05 for a total culture volume of 5 mL, and we
349 incubated this second pre-culture with shaking at 28°C for ~7 hours until it reached OD_{600} = 0.2.
350 For unstressed controls in Figure 1, we diluted an inoculum of the second pre-culture into fresh
351 liquid YPD at the volume required to attain OD_{600} = 0.2 for a total culture volume of 5 mL. 150 µL
352 of this experimental culture was transferred to each of 12 wells of a 96-well plate, sealed with a
353 Breathe-Easy membrane (Diversified Biotech), for incubation at 28°C and growth
354 measurements across a 24-hour timecourse in a Tecan M200 PRO plate reader (parameter
355 time point measurement, OD_{595}). For heat stress treatment in Figure 1, setup and growth
356 assays were as above except that incubation of the experimental culture was at 42°C. For
357 chemical stress treatments in Figure 1, setup and growth assays were as above except that for
358 the experimental culture, before transfer to the 96-well plate, ethanol (Koptec) was added to
359 attain a concentration of 6.5% (v/v); caffeine (Sigma-Aldrich) was added to attain 25 mM;
360 methane methylsulfonate (MMS; Thermo Scientific) was added to attain 0.03% (v/v); or
361 propidium iodide was added to attain 1mM, and growth timecourses were 48 to 72 hours. For
362 chemical stress treatments and growth measurements in Figures S1 and S2, for a given
363 species and condition, streaking and the first pre-culture were as above. We diluted an
364 inoculum of the first pre-culture into fresh liquid YPD at the volume required to attain OD_{600} = 0.2
365 for a total culture volume of 5 mL, with 1 M NaCl (Fisher), 3% glycerol (v/v; Fisher), 80 mM
366 $CaCl_2$ (Sigma), 250 mM rapamycin (Research Products International), 6 mM H_2O_2 (Fisher), or
367 16% (v/v) Triton X-100 (Fisher), followed by incubation and growth measurements in the 96-well
368 format as above.
369
370 **3. Assays of cell survival from *K. marxianus* and *K. lactis* stationary and log-phase liquid**
371 **culture**
372

373 For survival assays of cells in exponential phase in Figures 2 and S3, for *K. marxianus* Km31
374 and separately for *K. lactis* yLB72 we streaked a frozen stock onto yeast peptone dextrose
375 (YPD) agar plates, incubated for two days at 28°C, picked a single colony, inoculated into 5 mL
376 liquid YPD, and incubated with shaking at 28°C for 16-18 hours. We diluted, into fresh liquid
377 YPD, an inoculum of this first pre-culture at the volume required to attain OD_{600} = 0.05 for a total
378 culture volume of 5 mL, and we incubated this second pre-culture with shaking at 28°C for ~7
379 hours until it reached OD_{600} = 0.2. We took this second pre-culture directly into experimental
380 treatment as follows. For unstressed controls, serial dilutions of the culture were spotted onto
381 agar plates and incubated for 2 days at 28°C. For heat shock, 1 mL of the culture was incubated
382 in a heat block at 47°C for 30 minutes and then transferred to ice for 2 minutes, followed by
383 spotting and incubation as above.
384
385 For survival assays of cells in stationary phase in Figures 2 and S3, for *K. marxianus* Km31 and
386 separately for *K. lactis* yLB72 we streaked a frozen stock onto yeast peptone dextrose (YPD)
387 agar plates, incubated for two days at 28°C, picked a single colony, inoculated into 5 mL liquid
388 YPD, and incubated with shaking at 28°C for 96 hours. Heat shock and spotting (and
389 unstressed controls) were as above.
390
391 **4. Assays of cell survival from *K. marxianus* and *K. lactis* anaerobic and aerobic liquid**
392 **cultures**
393
394 As *Kluyveromyces* cannot tolerate strictly anaerobic conditions (Kiers et al. 1998; Ishtar Snoek
395 and Yde Steensma 2006; Merico et al. 2007; Dekker et al. 2021), for Figure S4 we used a
396 limited-oxygen experimental paradigm (Hoshida et al. 2020) as follows. For *K. marxianus* Km31
397 and separately for *K. lactis* yLB72, for each treatment we isolated a single colony on solid plates
398 and carried out first and second pre-cultures in liquid YPD at 28°C under aerobic conditions as

399 in Section 2. For unstressed controls in anaerobic conditions, we diluted an inoculum of the
400 second pre-culture into fresh liquid YPD at the volume required to attain OD_{600} = 0.05 for a total
401 culture volume of 5 mL. 150 µL of this experimental culture was transferred to each of 12 wells
402 of a 96-well plate, which was placed (with lid) into an AnaeroPouch (ThermoScientific) with a
403 AnaeroPouch - Anaero gas pack. The sealed pouch and plate were incubated for 72 hours at
404 28°C. Then we collected 30 µL of culture from each well and combined them all into one master
405 culture, 100 µL of which we serially diluted into spots on agar plates, followed by incubation for
406 2 days at 28°C. Aerobic controls were treated analogously except that the 96-well plate was
407 inoculated without a pouch. Viability assays of chemically stressed cells in anaerobic conditions
408 and aerobic controls were as above except that for the experimental culture, before transfer to
409 the 96-well plate, ethanol, caffeine, or MMS were added as in Section 2. Viability assays of
410 heat-treated cells in anaerobic conditions and aerobic controls were as above except that the
411 sealed pouch and plate were incubated for 24 hours at 39°C.
412
413 **5. Transcriptional profiling**
414
415 *5.1. Culture protocol and experimental design.* For a given replicate of RNA-seq of *K. marxianus*
416 Km31 and separately for *K. lactis* yLB72 in unstressed control conditions, we isolated a single
417 colony on solid plates and carried out first and second pre-cultures in liquid YPD at 28°C as in
418 Section 2. For unstressed controls, we diluted an inoculum of the second pre-culture into fresh
419 liquid YPD at the volume required to attain OD_{600} = 0.05 for a total culture volume of 50 mL. This
420 experimental culture was incubated at 28°C with shaking until it reached OD_{600} = 0.4-0.85. From
421 this culture we took a volume corresponding to 2 OD units, pelleted its cells by centrifugation for
422 5 minutes at 1,500 x g, and resuspended in 1 mL of RNAlater (ThermoFisher) followed by
423 storage at 4°C for up to 48 hours. For RNA-seq of heat-stressed cells, culture, pelleting, and
424 storage were as above except that the experimental culture was incubated at 37°C or 39°C; for

425 RNA-seq of cells under chemical stress, culture, pelleting, and storage were as above except
426 that the experimental culture was done in YPD with 22 mM caffeine, 0.01% MMS, or 3%
427 ethanol. For expression-based reannotation of *K. marxianus* Km31 and *K. lactis* yLB72 (see
428 below), we used the above pipeline for a preliminary expression profiling experiment which we
429 call the training batch: we cultured one biological replicate of Km31 and yLB72 at 37°C and one
430 of Km31 and yLB72 in unstressed control conditions. For comprehensive transcriptional profiling
431 analyses in Figure 3 and Figure S5, we used the above pipeline to generate cell pellets in two
432 production batches. In one batch we grew three biological replicates of Km31 and three of
433 yLB72 in 39°C heat stress, and three unstressed cultures of each species. In another batch we
434 grew three biological replicates of each species in caffeine, three in MMS, and three in ethanol,
435 and three unstressed controls.
436
437 *5.2. RNA isolation and sequencing*. For RNA isolation of a given sample pellet stored in
438 RNAlater (see Section 5.1), we centrifuged for 5 minutes at 1,500 x g, and after decanting of the
439 supernatant, we resuspended the pellet in 200 µL of nuclease-free water followed by addition of
440 750 µL of BashingBead buffer from the Direct-Zol RNA MicroPrep RNA kit (Zymo Research).
441 We then transferred our 950 µL of suspension into BashingBead tubes, vortexed for 20
442 minutes, and centrifuged for 5 minutes at 12,000 x g. We took ~600 µL of supernatant from the
443 top layer being careful not to touch the pellet, transferred to a new tube, and mixed with an
444 equal volume of 100% ethanol. This mixture was used as input to the rest of the kit, following
445 the manufacturer's instructions. RNA library preparation and sequencing of ~20M 150 bp
446 paired-end reads per sample was carried out by Novogene, Inc.
447
448 *5.3. Genome re-annotation and ortholog calls for* K. marxianus *and* K. lactis
449

450 To generate a reference genome for *K. marxianus* strain Km31, we proceeded as follows. We
451 first used bwa (bio-bwa.sourceforge.net) to map the RNA-seq reads from the *K. marxianus*
452 cultures of the training batch of our transcriptional profiling (see Section 5.1) to a published
453 genome of the DMB1 strain of *K. marxianus* (Suzuki et al. 2014). We then used these
454 mappings, requiring mapping qualities of >20 and mapped lengths >20, to call single-nucleotide
455 polymorphisms with BCFtools (Li 2011; Danecek et al. 2021) mpileup with the -I flag to ignore
456 indels, and filtering for exclusion of variants with: a depth of less than 10 reads, or more than
457 10% of reads with a mapping quality of 0 supporting the variant. At the called polymorphisms,
458 we computationally introduced Km31 alleles into the DMB1 genome sequence with the
459 consensus tool in BCFTools to generate a first-pass pseudogenome for Km31. Next, we used
460 this pseudogenome and the training batch RNA-seq reads as input into transcript assembly with
461 Trinity (Grabherr et al. 2011) and annotation with PASA (Haas et al. 2003), and we took the
462 output as our final annotated genome for *K. marxianus* Km31 (Supplementary Data 1). We
463 generated a reference genome for *K. lactis* strain yLB72 as above, except that we used RNA-
464 seq from the *K. lactis* training batch and the published genome of *K. lactis* strain NRRL Y-1140
465 (Dujon et al. 2004) (Supplementary Data 1). To call orthologous gene pairs between *K.*
466 *marxianus* Km31 and *K. lactis* yLB72, we used our final annotations from the two species as
467 input into OrthoFinder (Emms and Kelly 2019). We retained all one-to-one orthology calls, *i.e.*
468 cases in which a single annotated gene from *K. marxianus* was called orthologous to a single
469 gene from *K. lactis*. We also sought to valorize cases in which OrthoFinder, with the goal of
470 finding orthologs in species B of a gene $G_{1,A}$ in species A, did not distinguish between *n*
471 candidate orthologous genes $G_{1...n,B}$ from species B. We evaluated each candidate $G_{i,B}$ for
472 synteny: if the next gene up- or downstream of $G_{i,B}$ in the species B genome was called a one-
473 to-one ortholog of the up- or downstream genes of $G_{1,A}$ in the species A genome, we retained
474 this case as a bona fide ortholog pair. Our final set comprised 3578 ortholog pairs between *K.*

475 *marxianus* Km31 and *K. lactis* yLB72. *Saccharomyces* orthologs were taken from (Sorrells et al.
476 2018).
477
478 *5.4. Production-batch transcriptional profiling data analysis*
479
480 For comprehensive transcriptional profiling analyses in Figure 3 and Figure S5, we used our
481 RNA-seq data from cultures of *K. marxianus* Km31 and *K. lactis* yLB72 from production batches
482 (see Section 5.1) as follows. For the RNA-seq reads from a given replicate of a given species
483 (*K. marxianus* or *K. lactis*) and condition (39°C, ethanol, caffeine, MMS, or unstressed control),
484 we mapped to the respective pseudogenome (see Section 5.3) using the TopHat (Kim et al.
485 2013) aligner with default parameters, and we used these mappings as input into HTseq
486 (Anders et al. 2015) to generate raw read counts per gene. For a given species and stress, we
487 used the counts from all replicates from the stress, and the unstressed control from the
488 respective batch, as input into edgeR (Robinson et al. 2010) to generate normalized counts per
489 gene and replicate and a normalized $\log_2$ fold-change between stress and control conditions as
490 an average across replicates. Principal component analysis (PCA) of the RNA-seq data was
491 done using the PCA module of scikit-learn (Pedregosa et al. 2011). The edgeR normalized
492 counts for each gene in each treatment were scaled to unit variance with scikit-learn before
493 running the PCA. Principal component weights are listed in Supplementary Table 2.
494
495 To identify gene groups in which a given stress treatment evoked significant gene induction or
496 repression relative to the unstressed control, in *K. marxianus* Km31 and separately in *K. lactis*
497 yLB72 (Table S3), we carried out resampling-based enrichment tests as follows. We used
498 assignments of each gene to Gene Ontology terms for *K. lactis* from FungiDB
499 (https://fungidb.org/fungidb/app/record/dataset/NCBITAXON_284590), and we propagated
500 these assignments to orthologous genes in *K. marxianus* from Section 5.3. For a given species

501 and stress, for each Gene Ontology term in turn, for the *n* genes from the term with expression
502 observations, if $n > 5$ we collated the genes' $\log_2$ fold-changes in expression between stress
503 and the untreated control, each of which was a mean across replicates (from edgeR; see
504 above). We took the mean of these fold-changes across the genes of the term, m_{true}. We then
505 picked *n* random genes from the genome and tabulated their mean expression fold-change for
506 the respective species and stress, $m_{resample}$. Repeating the latter resampling 1000 times, we
507 calculated an empirical *p*-value reporting the enrichment of the magnitude of expression fold-
508 change as the percentage of resamples for which ($|m_{resample}| \geq (|m_{true}|$). We used the *p*-values for
509 all terms tested for the respective species and stress as input into multiple testing correction
510 with the Benjamini-Hochberg method.
511
512 **7. Molecular-evolution analyses**
513
514 *7.1. Gene duplication analysis*
515
516 As input to the gene family copy number analysis for Figure 4 and Table S4, genomes and
517 annotations were as follows. For *K. marxianus* and *K. lactis* we used the reference genomes
518 and annotations of Km31 and yLB72, respectively, from Section 5.3. For *K. wickerhamii* we
519 used the reference genome and annotation of UCD54-210 from (Sorrells et al. 2018). For the *K.*
520 *nonfermentans* strain NRRL Y-27343 (GenBank acc. no. GCA_030569915.1), the *K. aestuarii*
521 strain NRRL YB-4510 (GenBank acc. no. GCA_003707555.2), the *K. siamensis* strain CBS
522 10860 (GenBank acc. no. GCA_030579315.1), and the *K. dobzhanskii* strain NRRL Y-1974
523 (GenBank acc. no. GCA_003705805.3) we used YGAP (Proux-Wéra et al. 2012) to generate *de*
524 *novo* annotations from previously generated reference genomes (Supplementary Data 1). As
525 outgroup species we used the genomes and annotations for *Ashbya aceri* (GenBank acc. no.

526 GCA_000412225.2), *Candida albicans* (GenBank acc. no. GCA_000182965.3), *Eremothecium*
527 *cymbalariae* (GenBank acc. no. GCA_000235365.1), *Eremothecium gossypii* (GenBank acc.
528 no. GCA_000091025.4), *Eremothecium sinecaudum* (GenBank acc. no. GCA_001548555.1),
529 *Lachancea dasiensis* (GenBank acc. no. GCA_900074725.1), *Lachancea fantastica* (GenBank
530 acc. no. GCA_900074735.1), *Lachancea fermentatii* (GenBank acc. no. GCA_900074765.1),
531 *Lachancea lanzarotensis* (GenBank acc. no. GCA_000938715.1), *Lachancea meyersii*
532 (GenBank acc. no. GCA_900074715.1), *Lachancea mirantina* (GenBank acc. no.
533 GCA_900074745.1), *Lachancea nothofagi* (GenBank acc. no. GCA_900074755.1), *Lachancea*
534 *quebecensis* (GenBank acc. no. GCA_002900925.1), *Lachancea thermotolerans* (GenBank
535 acc. no. GCA_000142805.1), and *Saccharomyces cerevisiae* (GenBank acc. no.
536 GCA_000146045.2). We used proteomes for each species as input into OrthoFinder (Emms
537 and Kelly 2019) to find groups of orthologs. We then used the time-weighted tree and gene
538 family counts from OthoFinder as input into CAFE 5 (Mendes et al. 2021) to analyze gene
539 family contractions and expansions. Gene families that did not exist at the root were included.
540 We adjusted *p*-values with the Benjamini-Hochberg test to correct for multiple testing. We used
541 CafePlotter (https://github.com/moshi4/CafePlotter) to visualize significant gene family
542 expansions and contractions on the phylogenetic tree. We then assigned functions for each
543 gene family by running all proteins from all species for each gene family through InterProScan
544 (Jones et al. 2014). We took the most common InterPro term in the output of the scan as the
545 inferred function of the gene family in Figure 4.
546
547 *7.2. Phylogenetic analyses*
548
549 As input to phylogenetic analyses of one strain for each *Kluyveromyces* species for Figure 5
550 and Table S5, we used genomes and annotations for the *Kluyveromyces* genus excluding *K.*
551 *siamensis* as in Section 7.1. We called orthologs across these six species with OrthoFinder

552 (Emms and Kelly 2019), and for a given orthogroup we generated protein alignments using
553 MUSCLE (Edgar 2004). We then used this alignment and its nucleotide sequences as input into
554 PAL2NAL (Suyama et al. 2006) using the -nogap flag to remove gaps to generate a codon
555 alignment. For each orthogroup in turn, we ran the PAML (Yang 2007) codeml package run
556 using the branch-site model (Zhang et al. 2005) to evaluate the fit to the sequence data of a
557 model in which regions of the coding sequence are under positive selection for amino acid
558 variation along a focal (foreground) branch and not on any other branch from the tree (the
559 background branches). We used as input to PAML the codon alignment for the orthogroup and
560 the topology of the phylogenetic tree relating the species (as in Figure 1A), omitting *K.*
561 *siamensis*. The calculation proceeded in two steps. First, we fit a null model in which all
562 branches of the tree had the same protein evolutionary rate (PAML codeml run with options
563 verbose = 1, seqtype = 1, clock = 0, model = 2, NSsites = 2, fix_omega = 1, omega = 1),
564 yielding a protein evolutionary rate and log-likelihood of the data under the model. Next, we re-
565 fit using an alternative model in which the branch leading to *K. marxianus* was the foreground
566 (PAML codeml run with options verbose = 1, seqtype = 1, clock = 0, model = 2, NSsites = 2,
567 fix_omega = 0). Outputs were the codon sites assigned to the class with an excess of amino
568 acid changes in *K. marxianus* relative to silent changes, and their posterior probabilities under
569 Bayes Empirical Bayes inference (Yang et al. 2005); the codon sites assigned to classes with
570 relaxed and purifying selection, respectively, along the background lineages, and their posterior
571 probabilities; and the log-likelihood of the data under the model. The *p*-value assessing the
572 improvement in fit to the data of the alternative model relative to the null was calculated from a
573 likelihood ratio test (a χ^2 test on twice the difference of likelihoods with one degree of freedom).
574 *p*-values were corrected for multiple testing by the Benjamini/Hochberg method using the
575 multipletests function from statsmodels (Seabold and Perktold 2010). In Figure 5A we tabulated
576 all genes with adjusted $p < 0.05$ from this test that had more than 1% of codon sites assigned to
577 the class with an excess of amino acid changes in *K. marxianus* (proportion in class 2a > 0.01

578 in codeml). Separately, we repeated the above analyses using the branch leading to *K. lactis* as
579 the foreground. Conservation of sites called significant by PAML using *K. marxianus* as the
580 foreground were calculated using Biopython (Cock et al. 2009) to count the conservation of
581 each site among our sequenced *K. marxianus* strains (Table S1 and see Section 7.3).
582 *Saccharomyces* orthologs were taken from (Sorrells et al. 2018).
583
584 *7.3. Resequencing and analysis of* K. marxianus *and* K. lactis *population genomes*
585
586 For resequencing, we cultured, in 10 mL of liquid YPD, each of five *K. marxianus* strains of non-
587 domesticated provenance beside Km31, and two non-domesticated *K. lactis* strains beside
588 yLB72 (Table S1), and we isolated DNA from each with the Quick-DNA™ Fungal/Bacterial
589 Miniprep Kit (Zymo Research). Genomic DNA sequencing libraries were made and sequenced
590 on an Illumina NovaSeq 6000 by the UC Berkeley Vincent J. Coates Genomic Sequencing
591 Laboratory. We used Minimap2(Li 2018) to map the *K. marxianus* reads to our *K. marxianus*
592 Km31 reference genome (see Section 5.3). Separately, we mapped the reads from each of 41
593 *K. lactis* strains previously sequenced (Friedrich et al. 2023), and those from our two sequenced
594 *K. lactis* strains, to our *K. lactis* yLB72 reference genome. For each strain we sorted mappings
595 using the SAMtools fixmate and markdup functions, and we called single-nucleotide
596 polymorphisms and generated a pseudogenome as in Section 5.3, except the variant filtering
597 included only a quality score (QUAL) threshold of above 20. Called variants are provided in
598 Supplementary Data 2. For each pair of orthologous genes between *K. marxianus* and *K. lactis*
599 from section 5.3, we aligned the nucleotide sequences of the coding regions from all strains of
600 the two species using MUSCLE (Edgar 2004). We then used this alignment and its associated
601 nucleotide sequences as input into PAL2NAL (Suyama et al. 2006) using the -nogap flag to
602 remove gaps to generate a codon alignment.
603

604 For Figure 5D and 5E and Table S6 and S7, we carried out per-gene McDonald-Kreitman
605 analyses (McDonald and Kreitman 1991) on the PAL2NAL alignments using the mktest function
606 from the CodonAlignment class from Biopython. *p*-values were corrected for multiple testing by
607 the Benjamini/Hochberg method using the multipletests function from statsmodels (Seabold and
608 Perktold 2010). For a given gene, this approach tabulates the rates of non-synonymous
609 polymorphism within species (P_n), synonymous polymorphism (P_s), non-synonymous
610 divergence between species (D_n), and synonymous divergence (D_s), and the neutrality index NI
611 = $(P_n/P_s)/(D_n/D_s)$. Because low counts of amino acid polymorphism can drive significance in the
612 McDonald-Kreitman framework, we developed a filtering scheme in which we tabulated the
613 average a and standard deviation σ of P_n/P_s across all genes for which we had sequence data.
614 We only retained for Figure 5D those genes with $P_n/P_s > a - 1.5\sigma$.
615
616 For genomic analyses of P_n/P_s and D_n/D_s in the McDonald-Kreitman framework in Figure 5E and
617 Table S7, we used Gene Ontology term assignments and resampling (using 10,000 resamples)
618 as in Section 5.4, except that we only included Gene Ontology terms containing 10 genes or
619 more after filtering, and for P_n/P_s we tested for enrichment of low values of the metric across a
620 given term. Genomic analyses on NI proceeded as above, except on the filtered dataset of
621 genes with $P_n/P_s > a - 1.5\sigma$ as explained above and testing for enrichment of low values of the
622 metric across the term. *p*-values were corrected for multiple testing by the Benjamini/Hochberg
623 method using the multipletests function from statsmodels (Seabold and Perktold 2010).
624

625 ## Acknowledgements

626

627 The authors thank Jasmine Goodman, Sasha Padilla, Catherine Rha, and Johnson Jun-Ting
628 Wang for experimental support; Adam Arkin for his generosity with computational resources;

Jamie Cate and Alexander Johnson for strains; and Jeremy Roop and members of the Brem lab for helpful discussions. This work was supported by NIH 2R01GM120430 to R.B.B. and an NSF Graduate Research Fellowship to A.D.

Data Availability

RNA-seq reads and normalized read counts will be made available in the Gene Expression Omnibus, and genome sequencing reads in the Sequence Read Archive, upon publication.

637 # Figure captions

638

639 **Figure 1. *K. marxianus* shows a unique heat and chemical stress resistance phenotype.**
640 **A,** Phylogenetic tree of the *Kluyveromyces* genus with marine and terrestrial species delineated
641 (Am-In et al. 2008; Varela, Puricelli, Ortiz-Merino, et al. 2019). Branch lengths are not to scale.
642 **B,** Each panel reports the results of a growth profiling experiment across the genus in the
643 indicated condition. In a given panel, each trace reports a growth timecourse for the indicated
644 species, with the solid line displaying the mean across technical replicates and the faint outline
645 showing the standard error. OD, optical density.

646

647 **Figure 2. The unique heat tolerance of *K. marxianus* only manifests during its growing**
648 **stage. A**, Experimental workflow created with BioRender.com. From left, inoculum from a
649 starting liquid culture is introduced into fresh liquid medium for regrowth at 28°C and then
650 sampled at nutrient exhaustion (Stationary) or in log phase (Exponential). The liquid sample is
651 subjected to a heat shock of 47°C for 30 minutes and then spotted as serial dilutions onto solid
652 medium. Viable cells after heat shock grow into colonies after incubation at 28°C. **B**-**C**, Each
653 row reports viability after heat shock of a sample of the indicated species in the indicated liquid
654 growth phase. See Figure S3 for unshocked controls.

655

656 **Figure 3. Overall regulatory volatility by *K. lactis* relative to *K. marxianus* and stress-**
657 **evoked repression of glycolysis in both species. A,** Shown are results of principal
658 component (PC) analysis of *K. lactis* and *K. marxianus* transcriptomes across treatments. Each
659 point reports the values of the top two PCs for one replicate transcriptome of the indicated
660 species subjected to the indicated stress treatment or a 28°C unstressed control. In a given axis
661 label, the value in parentheses reports the proportion of variance across the set of

662 transcriptomes explained by the indicated PC. **B,C,** Each panel reports the effect of stress on
663 expression of genes of the indicated Gene Ontology term in *K. marxianus* and *K. lactis*. In a
664 given panel, each cell reports, for the indicated gene, the $\log_2$ fold-change between expression
665 levels in the indicated stress treatment and an unstressed control.
666
667 **Figure 4. Gene family expansions in *K. marxianus*. A,** The color of each cell reports the
668 change in copy number, in the indicated genome of a *Kluyveromyces* or related species relative
669 to its parent node in the phylogeny, of a gene family (orthogroup, OG) with significant expansion
670 or contraction in the *K. marxianus* genome as inferred with CAFE 5. Results for all analyzed
671 orthogroups are reported in Table S4. **B,** Each panel reports the positions of genes in the *K.*
672 *marxianus* genome in one indicated gene family from **A** with annotations as major facilitator
673 superfamily (MFS) transporters. In a given panel, each row represents a chromosome in *K.*
674 *marxianus*, and each arrow reports a gene of the respective family and its transcriptional
675 direction (not to scale).
676
677 **Figure 5. Phylogenetic analysis reveals accelerated amino acid evolution at hundreds of**
678 **genes in *K. marxianus*. A,** Each bar shows the number of genes with significant evidence for
679 positive selection on the indicated lineage (adjusted $p > 0.05$) in a phylogenetic PAML branch-
680 site test. **B,** The set of bars of each color reports the distribution, across the genes from **A**, of
681 the number of sites per gene inferred by PAML to be targets of positive selection on the
682 indicated lineage. **C,** Amino acid alignment of an example region of a gene, KLLA0B02607g (*S.*
683 *cerevisiae* ortholog FAT3/*YKL187C*), with evidence for positive selection from the PAML test in
684 **A** and from the McDonald-Kreitman test of **D**. Highlighted columns show inferred targets of
685 positive selection from the PAML test in **B**. Highlighting indicates the amino acid's BLOSUM 62
686 matrix score with respect to the consensus amino acid. **D,** Each row reports a gene with
687 evidence for adaptive protein variation between *K. marxianus* and *K. lactis* according to the

688 McDonald-Kreitman test (adjusted $p < 0.05$). NI, neutrality index. Genes highlighted in green
689 also showed significant evidence for positive selection on the branch leading to *K. marxianus* in
690 phylogenetic tests (PAML branch-site $p < 0.05$; Table S5). Not shown here are genes with *p*-
691 value < 0.05 as a product of exceptionally high or low amino acid polymorphism; see Methods
692 and Table S5. **E**, The *x*-axis reports mean NI across a group of genes. The blue trace reports
693 results from a resampling test for enrichment of low NI across the genes of the transmembrane
694 transport Gene Ontology term: the *y*-axis reports the frequency, among random sets of genes,
695 of the mean NI given on the *x*, and the true mean NI of genes of the transmembrane transport
696 term is indicated by the red dashed line, corresponding to corrected empirical $p = 0.0132$ (Table
697 S5).

698 ## Supplementary figure captions

699

700 **Figure S1. Stresses in which *K. marxianus* does not exhibit unique stress tolerance**
701 **relative to the rest of the genus.** Data and symbols are as in Figure 1B of the main text except
702 that each panel reports growth measurements in the indicated stress.

703

704 **Figure S2. *K. marxianus* strains mirror each other in heat tolerance.** Data and symbols are
705 as in Figure 1B of the main text except that in a given panel, each trace reports growth
706 measurements of the indicated wild *K. marxianus* strain.

707

708 **Figure S3. No detectable difference between *Kluyveromyces* species with respect to cell**
709 **viability from unstressed liquid cultures.** Data and symbols are as in Figure 2 of the main
710 text except that liquid cultures were not subjected to heat shock before plating.

711

712 **Figure S4. Genetic or culture-based block of respiration compromises stress tolerance in**
713 ***K. marxianus* and *K. lactis*. A, B,** Each panel reports cell viability assays after culture in liquid
714 medium with stress, or an unstressed control, of *K. marxianus* or *K. lactis* in normoxia or
715 hypoxia. In a given panel, each row shows serial dilutions of the liquid culture spotted onto solid
716 medium and incubated without stress at 28°C.

717

718 **Figure S5. Stress-evoked changes in *K. lactis*, but not *K. marxianus,* dominate**
719 **transcriptome comparisons between the two. A,** Shown are results from the principal
720 component (PC) analysis of *K. lactis* and *K. marxianus* transcriptomes across treatments in
721 Figure 3A of the main text. Each panel reports expression across genotypes and treatments of
722 the three genes contributing most to the indicated PC. In a given panel, each bar reports

723 expression, as a mean across replicates, of the indicated gene in the indicated species and
724 treatment. The full set of gene weights is listed in Supplementary Table 2. **B, C**, Data are as in
725 Figure 3B,C of the main text except that genes of the indicated Gene Ontology term are
726 reported.

727 ## Supplementary table captions

728

729 **Table S1. Strains used in this work.** List of strains used for phenotypic assays and population
730 analyses. All species shown are haploid.

731

732 **Table S2. Normalized RNA-seq read counts, major principal component weights, and**
733 **stress effects.** Each row reports RNA-seq results and analysis for a pair of orthologous *K.*
734 *lactis* and *K. marxianus* genes. Columns B and C report respective weighting of the gene in the
735 first and second components of a principal component analysis. In columns D-S, each column
736 reports normalized RNA-seq read counts as an average across replicates from cultures of *K.*
737 *lactis* or *K. marxianus* in unstressed control conditions (LC and MC, respectively) or in stress
738 treatments (LT or MT, respectively), where stresses are 39°C, ethanol (ET), MMS, or caffeine
739 (CAF). The last eight columns report the effect of stress on expression; each reports the
740 average across replicates of $\log_2$(treatment/control) for each species in the indicated treatment.

741

742 **Table S3. Gene Ontology terms with stress-responsive expression.** Each tab reports the
743 results of tests of Gene Ontology terms for enrichment of expression change between the
744 indicated stress and unstressed controls in the indicated species. In a given tab, each row
745 reports results for the indicated Gene Ontology term: the average, across genes and replicates,
746 of $\log_2$(treatment/control), and the multiple testing-corrected significance of the enrichment for
747 high magnitude of this value relative to a resampling null.

748

749 **Table S4. Gene family analysis for expansion/contraction analysis.** Tab 1 reports the
750 results from the CAFE analysis for gene family copy number change. Column B denotes the *p*-
751 value from CAFE for a change in copy number in *K. marxianus* of the orthogroup indicated in

752 Column A. Column C shows the *p*-value after adjustment for multiple testing. Tab 2 shows the
753 genes called in each orthogroup for all species analyzed in CAFE.
754
755 **Table S5. Phylogenetic tests for evidence of positive selection.** In a given row in Tab 1,
756 columns C-I report the results of PAML branch-site tests package for positive selection in *K.*
757 *marxianus* on the indicated orthogroup (see Tab 2 for full orthogroup membership). Columns C-
758 D report the significance of the improvement in fit to an alternative model stipulating positive
759 selection relative to a null without positive selection, as a nominal *p*-value and after adjustment
760 for multiple testing. Columns E and G report the number and proportion, respectively, of sites
761 inferred to be under positive selection in the alternative model. Column F reports the average
762 conservation of the latter across *K. marxianus* population isolates from Supplementary Data 2.
763 Columns H and I report inferred protein evolutionary rates for the branch leading to *K.*
764 *marxianus* (foreground) and for all other branches of the tree (background), respectively.
765 Columns J-O report the analogous results from tests for positive selection in *K. lactis*. Columns
766 P-Q list the gene name and function for the *S. cerevisiae* ortholog. Green highlighted cells
767 indicate cases of $p < 0.05$. Tab 2 shows the genes in each orthogroup analyzed in PAML, with
768 each row reporting the genes of the indicated species representing the members of one inferred
769 orthogroup.
770
771 **Table S6. McDonald-Kreitman tests for evidence of positive selection.** Each row reports
772 results from a McDonald-Kreitman test on the indicated pair of orthologous genes between *K.*
773 *lactis* and *K. marxianus*. D_s, number of divergent synonymous changes between species; D_n,
774 number of divergent non-synonymous changes between species; P_s, number of polymorphic
775 synonymous changes within species; P_n, number of polymorphic non-synonymous changes
776 within species; NI, neutrality index (see Methods); *p* and *p*-value, nominal and multiple testing-

777 corrected McDonald-Kreitman significances, respectively. Columns L-M list the gene name and
778 function for the *S. cerevisiae* ortholog. Green highlighted cells indicate cases of $p < 0.05$.
779
780 **Table S7. Gene Ontology terms with evidence for positive or purifying selection.** Each tab
781 reports the results of tests of Gene Ontology terms for enrichment of extreme values of one
782 sequence-based metric of selection. These tests used the gene-based variation data from Table
783 S5 and were, respectively, for low values of the neutrality index, NI, indicative of positive
784 selection; for low values of P_s/P_n, indicative of purifying selection; and for high values of the
785 protein evolutionary rate D_s/D_n, indicative of positive selection. In a given tab, each row reports
786 results for the indicated Gene Ontology term: the average across genes of the indicated metric,
787 and the nominal and the multiple testing-corrected significance of the enrichment for extreme
788 magnitude of this value relative to a resampling null (p-value and p_adj, respectively).

789 # References

790

791 Abrams MB, Chuong JN, AlZaben F, Dubin CA, Skerker JM, Brem RB. 2022. Barcoded
792 reciprocal hemizygosity analysis via sequencing illuminates the complex genetic basis of
793 yeast thermotolerance. *G3 Genes|Genomes|Genetics* 12:jkab412.

794 Al-Qaysi SAS, Abdullah NM, Jaffer MR, Abbas ZA. 2021. Biological Control of Phytopathogenic
795 Fungi by Kluyveromyces marxianus and Torulaspora delbrueckii Isolated from Iraqi Date
796 Vinegar. *J Pure Appl Microbiol.* 15:300–311.

797 AlZaben F, Chuong JN, Abrams MB, Brem RB. 2021. Joint effects of genes underlying a
798 temperature specialization tradeoff in yeast. *PLOS Genetics* 17:e1009793.

799 Am-In S, Yongmanitchai W, Limtong S. 2008. Kluyveromyces siamensis sp. nov., an
800 ascomycetous yeast isolated from water in a mangrove forest in Ranong Province,
801 Thailand. *FEMS Yeast Research* 8:823–828.

802 Anders S, Pyl PT, Huber W. 2015. HTSeq—a Python framework to work with high-throughput
803 sequencing data. *Bioinformatics* 31:166–169.

804 Anderson PJ, McNeil K, Watson K. 1986. High-Efficiency Carbohydrate Fermentation to Ethanol
805 at Temperatures above 40°C by Kluyveromyces marxianus var. marxianus Isolated from
806 Sugar Mills. *Applied and Environmental Microbiology* 51:1314–1320.

807 Avchar R, Lanjekar V, Dhakephalkar PK, Dagar SS, Baghela A. 2022. Compost as an untapped
808 niche for thermotolerant yeasts capable of high-temperature ethanol production. *Letters*
809 *in Applied Microbiology* 74:109–121.

810 Baptista M, Cunha JT, Domingues L. 2021. Establishment of Kluyveromyces marxianus as a
811 Microbial Cell Factory for Lignocellulosic Processes: Production of High Value Furan
812 Derivatives. *Journal of Fungi* 7:1047.

813 Baptista M, Domingues L. 2022. Kluyveromyces marxianus as a microbial cell factory for
814 lignocellulosic biomass valorisation. *Biotechnology Advances* 60:108027.

815 Bell G. 2009. The Oligogenic View of Adaptation. *Cold Spring Harb Symp Quant Biol* 74:139–
816 144.

817 Blum M, Chang H-Y, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, Nuka G, Paysan-
818 Lafosse T, Qureshi M, Raj S, et al. 2021. The InterPro protein families and domains
819 database: 20 years on. *Nucleic Acids Research* 49:D344–D354.

820 Boyle EA, Li YI, Pritchard JK. 2017. An Expanded View of Complex Traits: From Polygenic to
821 Omnigenic. *Cell* 169:1177–1186.

822 Burton WG. 1982. Post-harvest Physiology of Food Crops. Longman

823 Castoe TA, de Koning APJ, Hall KT, Card DC, Schield DR, Fujita MK, Ruggiero RP, Degner JF,
824 Daza JM, Gu W, et al. 2013. The Burmese python genome reveals the molecular basis
825 for extreme adaptation in snakes. *Proceedings of the National Academy of Sciences*
826 110:20645–20650.

827 Cernak P, Estrela R, Poddar S, Skerker JM, Cheng Y-F, Carlson AK, Chen B, Glynn VM, Furlan
828 M, Ryan OW, et al. 2018. Engineering Kluyveromyces marxianus as a Robust Synthetic
829 Biology Platform Host. *mBio* 9:10.1128/mbio.01410-18.

830 Chacón-Vargas K, Torres J, Giles-Gómez M, Escalante A, Gibbons JG. 2020. Genomic profiling
831 of bacterial and fungal communities and their predictive functionality during pulque
832 fermentation by whole-genome shotgun sequencing. *Sci Rep* 10:15115.

833 Chikina M, Robinson JD, Clark NL. 2016. Hundreds of Genes Experienced Convergent Shifts in
834 Selective Pressure in Marine Mammals. *Molecular Biology and Evolution* 33:2182–2192.

835 Cock PJA, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff
836 F, Wilczynski B, et al. 2009. Biopython: freely available Python tools for computational
837 molecular biology and bioinformatics. *Bioinformatics* 25:1422–1423.

838 Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T,
839 McCarthy SA, Davies RM, et al. 2021. Twelve years of SAMtools and BCFtools.
840 *GigaScience* 10:giab008.

841 Dekker WJC, Ortiz-Merino RA, Kaljouw A, Battjes J, Wiering FW, Mooiman C, Torre P de la,
842 Pronk JT. 2021. Engineering the thermotolerant industrial yeast Kluyveromyces
843 marxianus for anaerobic growth. *Metabolic Engineering* 67:347–364.

844 Dujon B, Sherman D, Fischer G, Durrens P, Casaregola S, Lafontaine I, de Montigny J, Marck
845 C, Neuvéglise C, Talla E, et al. 2004. Genome evolution in yeasts. *Nature* 430:35–44.

846 Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high
847 throughput. *Nucleic Acids Research* 32:1792–1797.

848 Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative
849 genomics. *Genome Biology* 20:238.

850 Falcão Moreira R, Fernandes PA, Moradas-Ferreira P. 1998. Kluyveromyces marxianus
851 flocculence and growth at high temperature is dependent on the presence of the protein
852 p37. *Microbiology* 144:681–688.

853 Ferenci T. 2016. Trade-off Mechanisms Shaping the Diversity of Bacteria. *Trends Microbiol*
854 24:209–223.

855 Flores-Cosío G, García-Béjar JA, Sandoval-Nuñez D, Amaya-Delgado L. 2024. Stress
856 response and adaptation mechanisms in *Kluyveromyces marxianus*. In: Gadd GM,
857 Sariaslani S, editors. Advances in Applied Microbiology. Vol. 126. Advances in Applied
858 Microbiology. Academic Press. p. 27–62. Available from:
859 https://www.sciencedirect.com/science/article/pii/S0065216424000030

860 Friedrich A, Gounot J-S, Tsouris A, Bleykasten C, Freel K, Caradec C, Schacherer J. 2023.
861 Contrasting Genomic Evolution Between Domesticated and Wild Kluyveromyces lactis
862 Yeast Populations. *Genome Biology and Evolution* 15:evad004.

863 Fu X, Li P, Zhang L, Li S. 2019. Understanding the stress responses of Kluyveromyces
864 marxianus after an arrest during high-temperature ethanol fermentation based on
865 integration of RNA-Seq and metabolite data. *Appl Microbiol Biotechnol* 103:2715–2729.

866 Gethins L, Rea MC, Stanton C, Ross RP, Kilcawley K, O'Sullivan M, Crotty S, Morrissey JP.
867 2016. Acquisition of the yeast Kluyveromyces marxianus from unpasteurised milk by a
868 kefir grain enhances kefir quality. *FEMS Microbiology Letters* 363:fnw165.

869 Goodspeed D, Liu JD, Chehab EW, Sheng Z, Francisco M, Kliebenstein DJ, Braam J. 2013.
870 Postharvest Circadian Entrainment Enhances Crop Pest Resistance and Phytochemical
871 Cycling. *Current Biology* 23:1235–1241.

872 Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L,
873 Raychowdhury R, Zeng Q, et al. 2011. Full-length transcriptome assembly from RNA-
874 Seq data without a reference genome. *Nat Biotechnol* 29:644–652.

875 Haas BJ, Delcher AL, Mount SM, Wortman JR, Smith Jr RK, Hannick LI, Maiti R, Ronning CM,
876 Rusch DB, Town CD, et al. 2003. Improving the Arabidopsis genome annotation using
877 maximal transcript alignment assemblies. *Nucleic Acids Research* 31:5654–5666.

878 Hayward LK, Sella G. 2022. Polygenic adaptation after a sudden change in environment.Coop
879 G, Perry GH, Ralph PL, Martin G, editors. *eLife* 11:e66697.

880 Hoshida H, Kagawa S, Ogami K, Akada R. 2020. Anoxia-induced mitophagy in the yeast
881 Kluyveromyces marxianus. *FEMS Yeast Research* 20:foaa057.

882 Illarionov A, Lahtvee P-J, Kumar R. 2021. Potassium and Sodium Salt Stress Characterization
883 in the Yeasts Saccharomyces cerevisiae, Kluyveromyces marxianus, and Rhodotorula
884 toruloides. *Applied and Environmental Microbiology* 87:e03100-20.

885 Ishtar Snoek IS, Yde Steensma H. 2006. Why does Kluyveromyces lactis not grow under
886 anaerobic conditions? Comparison of essential anaerobic genes of Saccharomyces
887 cerevisiae with the Kluyveromyces lactis genome. *FEMS Yeast Research* 6:393–403.

888 Jain K, Stephan W. 2017. Modes of Rapid Polygenic Adaptation. *Molecular Biology and*
889 *Evolution* 34:3169–3175.

890 Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A,
891 Nuka G, et al. 2014. InterProScan 5: genome-scale protein function classification.
892 *Bioinformatics* 30:1236–1240.

893 Kiers J, Zeeman A-M, Luttik M, Thiele C, Castrillo JI, Steensma HY, Van Dijken JP, Pronk JT.
894 1998. Regulation of alcoholic fermentation in batch and chemostat cultures of
895 Kluyveromyces lactis CBS 2359. *Yeast* 14:459–469.

896 Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate
897 alignment of transcriptomes in the presence of insertions, deletions and gene fusions.
898 *Genome Biology* 14:R36.

899 Kolora SRR, Owens GL, Vazquez JM, Stubbs A, Chatla K, Jainese C, Seeto K, McCrea M,
900 Sandel MW, Vianna JA, et al. 2021. Origins and evolution of extreme life span in Pacific
901 Ocean rockfishes. *Science* 374:842–847.

902 Kosaka T, Tsuzuno T, Nishida S, Pattanakittivorakul S, Murata M, Miyakawa I, Lertwattanasakul
903 N, Limtong S, Yamada M. 2022. Distinct Metabolic Flow in Response to Temperature in
904 Thermotolerant Kluyveromyces marxianus. *Applied and Environmental Microbiology*
905 88:e02006-21.

906 Kurtzman C, Fell JW, Boekhout T. 2011. The Yeasts: A Taxonomic Study. Elsevier

907 Lane MM, Burke N, Karreman R, Wolfe KH, O'Byrne CP, Morrissey JP. 2011. Physiological and
908 metabolic diversity in the yeast Kluyveromyces marxianus. *Antonie van Leeuwenhoek*
909 100:507–519.

910 Lane MM, Morrissey JP. 2010. Kluyveromyces marxianus: A yeast emerging from its sister's
911 shadow. *Fungal Biology Reviews* 24:17–26.

912 Lappe-Oliveras P, Avitia M, Sánchez-Robledo SD, Castillo-Plata AK, Pedraza L, Baquerizo G,
913 Le Borgne S. 2023. Genotypic and Phenotypic Diversity of Kluyveromyces marxianus
914 Isolates Obtained from the Elaboration Process of Two Traditional Mexican Alcoholic
915 Beverages Derived from Agave: Pulque and Henequen (Agave fourcroydes) Mezcal.
916 *Journal of Fungi* 9:795.

917 Lertwattanasakul N, Suprayogi, Murata M, Rodrussamee N, Limtong S, Kosaka T, Yamada M.
918 2013. Essentiality of respiratory activity for pentose utilization in thermotolerant yeast
919 Kluyveromyces marxianus DMKU 3-1042. *Antonie van Leeuwenhoek* 103:933–945.

920 Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping
921 and population genetical parameter estimation from sequencing data. *Bioinformatics*
922 27:2987–2993.

923 Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34:3094–
924 3100.

925 Linder T. 2018. Evaluation of the chitin-binding dye Congo red as a selection agent for the
926 isolation, classification, and enumeration of ascomycete yeasts. *Arch Microbiol* 200:671–
927 675.

928 Madeira-Jr JV, Gombert AK. 2018. Towards high-temperature fuel ethanol production using
929 Kluyveromyces marxianus: On the search for plug-in strains for the Brazilian sugarcane-
930 based biorefinery. *Biomass and Bioenergy* 119:217–228.

931 Masly JP, Azom MG. 2022. Molecular divergence with major morphological consequences:
932 development and evolution of organ size and shape. *Essays in Biochemistry* 66:707–
933 716.

934 McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the Adh locus in Drosophila.
935 *Nature* 351:652–654.

936 Mejía-Barajas JA, Martínez-Mora JA, Salgado-Garciglia R, Noriega-Cisneros R, Ortiz-Avila O,
937 Cortés-Rojo C, Saavedra-Molina A. 2017. Electron transport chain in a thermotolerant
938 yeast. *J Bioenerg Biomembr* 49:195–203.

939 Mendes FK, Vanderpool D, Fulton B, Hahn MW. 2021. CAFE 5 models variation in evolutionary
940 rates among gene families. *Bioinformatics* 36:5516–5518.

941 Merico A, Sulo P, Piškur J, Compagno C. 2007. Fermentative lifestyle in yeasts belonging to the
942 Saccharomyces complex. *The FEBS Journal* 274:976–989.

943 Montini N, Doughty TW, Domenzain I, Fenton DA, Baranov PV, Harrington R, Nielsen J,
944 Siewers V, Morrissey JP. 2022. Identification of a novel gene required for competitive
945 growth at high temperature in the thermotolerant yeast Kluyveromyces marxianus.
946 *Microbiology* 168:001148.

947 Nespolo RF, Solano-Iguaran JJ, Paleo-López R, Quintero-Galvis JF, Cubillos FA, Bozinovic F.
948 2020. Performance, genomic rearrangements, and signatures of adaptive evolution:
949 Lessons from fermentative yeasts. *Ecology and Evolution* 10:5240–5250.

950 Ortiz-Merino RA, Varela JA, Coughlan AY, Hoshida H, da Silveira WB, Wilde C, Kuijpers NGA,
951 Geertman J-M, Wolfe KH, Morrissey JP. 2018. Ploidy Variation in Kluyveromyces
952 marxianus Separates Dairy and Non-dairy Isolates. *Frontiers in Genetics* [Internet] 9.
953 Available from: https://www.frontiersin.org/articles/10.3389/fgene.2018.00094

954 Papouskova K, Akinola J, Ruiz-Castilla FJ, Morrissey JP, Ramos J, Sychrova H. 2024. The
955 superior growth of Kluyveromyces marxianus at very low potassium concentrations is
956 enabled by the high-affinity potassium transporter Hak1. *FEMS Yeast Research*
957 24:foae031.

958 Pease JB, Haak DC, Hahn MW, Moyle LC. 2016. Phylogenomics Reveals Three Sources of
959 Adaptive Variation during a Rapid Radiation. *PLOS Biology* 14:e1002379.

960 Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer
961 P, Weiss R, Dubourg V, et al. 2011. Scikit-learn: Machine Learning in Python. *Journal of*
962 *Machine Learning Research* 12:2825–2830.

963 Prado MR, Blandón LM, Vandenberghe LPS, Rodrigues C, Castro GR, Thomaz-Soccol V,
964 Soccol CR. 2015. Milk kefir: composition, microbial cultures, biological activities, and
965 related products. *Frontiers in Microbiology* [Internet] 6. Available from:
966 https://www.frontiersin.org/articles/10.3389/fmicb.2015.01177

967 Proux-Wéra E, Armisén D, Byrne KP, Wolfe KH. 2012. A pipeline for automated annotation of
968 yeast genome sequences by a conserved-synteny approach. *BMC Bioinformatics*
969 13:237.

970 Rainey TJ, Covey G. 2016. Pulp and paper production from sugarcane bagasse. In: Sugarcane-
971 Based Biofuels and Bioproducts. John Wiley & Sons, Ltd. p. 259–280. Available from:
972 https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118719862.ch10

973 Rij NJWK. 2013. The Yeasts: A Taxonomic Study. Elsevier

974 Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential
975 expression analysis of digital gene expression data. *Bioinformatics* 26:139–140.

976 Rockman MV. 2012. THE QTN PROGRAM AND THE ALLELES THAT MATTER FOR
977 EVOLUTION: ALL THAT'S GOLD DOES NOT GLITTER. *Evolution* 66:1–17.

978 Saini P, Beniwal A, Vij S. 2017. Physiological response of *Kluyveromyces marxianus* during
979 oxidative and osmotic stress. *Process Biochemistry* 56:21–29.

980 Seabold S, Perktold J. 2010. Statsmodels: Econometric and Statistical Modeling with Python.
981 In: Austin, Texas. p. 92–96. Available from:
982 https://conference.scipy.org/proceedings/scipy2010/seabold.html

983 Shen X-X, Opulente DA, Kominek J, Zhou X, Steenwyk JL, Buh KV, Haase MAB, Wisecaver
984 JH, Wang M, Doering DT, et al. 2018. Tempo and Mode of Genome Evolution in the
985 Budding Yeast Subphylum. *Cell* 175:1533-1545.e20.

986 Shimada TL, Takano Y, Hara-Nishimura I. 2015. Oil body-mediated defense against fungi: From
987 tissues to ecology. *Plant Signaling & Behavior* 10:e989036.

988 Smith SD, Pennell MW, Dunn CW, Edwards SV. 2020. Phylogenetics is the New Genetics (for
989 Most of Biodiversity). *Trends in Ecology & Evolution* 35:415–425.

990 Sorrells TR, Booth LN, Tuch BB, Johnson AD. 2015. Intersecting transcription networks
991 constrain gene regulatory evolution. *Nature* 523:361–365.

992 Sorrells TR, Johnson AN, Howard CJ, Britton CS, Fowler KR, Feigerle JT, Weil PA, Johnson
993 AD. 2018. Intrinsic cooperativity potentiates parallel cis-regulatory evolution. *eLife*
994 7:e37563.

995 Spurley WJ, Fisher KJ, Langdon QK, Buh KV, Jarzyna M, Haase MAB, Sylvester K, Moriarty
996 RV, Rodriguez D, Sheddan A, et al. 2022. Substrate, temperature, and geographical
997 patterns among nearly 2000 natural yeast isolates. *Yeast* 39:55–68.

998 Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence
999 alignments into the corresponding codon alignments. *Nucleic Acids Research* 34:W609–
1000 W612.

1001 Suzuki T, Hoshino T, Matsushika A. 2014. Draft Genome Sequence of Kluyveromyces
1002 marxianus Strain DMB1, Isolated from Sugarcane Bagasse Hydrolysate. *Genome*
1003 *Announcements* 2:10.1128/genomea.00733-14.

1004 Varela JA, Puricelli M, Montini N, Morrissey JP. 2019. Expansion and Diversification of MFS
1005 Transporters in Kluyveromyces marxianus. *Front. Microbiol.* [Internet] 9. Available from:
1006 https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2018.03330/full

1007 Varela JA, Puricelli M, Ortiz-Merino RA, Giacomobono R, Braun-Galleani S, Wolfe KH,
1008 Morrissey JP. 2019. Origin of Lactose Fermentation in Kluyveromyces lactis by

1009 Interspecies Transfer of a Neo-functionalized Gene Cluster during Domestication.
1010 *Current Biology* 29:4284-4290.e2.

1011 Verma S, Awasthi MK, Liu T, Awasthi SK, Syed A, Bahkali AH, Verma M, Zhang Z. 2023.
1012 Influence of biochar on succession of fungal communities during food waste composting.
1013 *Bioresource Technology* 385:129437.

1014 Weiss CV, Brem RB. 2019. Dissecting Trait Variation across Species Barriers. *Trends in*
1015 *Ecology & Evolution* 34:1131–1136.

1016 Weiss CV, Roop JI, Hackley RK, Chuong JN, Grigoriev IV, Arkin AP, Skerker JM, Brem RB.
1017 2018. Genetic dissection of interspecific differences in yeast thermotolerance. *Nat Genet*
1018 50:1501–1504.

1019 Williams DL, Schückel J, Vivier MA, Buffetto F, Zietsman AJJ. 2019. Grape pomace
1020 fermentation and cell wall degradation by Kluyveromyces marxianus Y885. *Biochemical*
1021 *Engineering Journal* 150:107282.

1022 Woodard SH, Fischman BJ, Venkat A, Hudson ME, Varala K, Cameron SA, Clark AG,
1023 Robinson GE. 2011. Genes involved in convergent evolution of eusociality in bees.
1024 *Proceedings of the National Academy of Sciences* 108:7472–7477.

1025 Wu D, Wang D, Hong J. 2020. Effect of a Novel Alpha/Beta Hydrolase Domain Protein on
1026 Tolerance of K. marxianus to Lignocellulosic Biomass Derived Inhibitors. *Frontiers in*
1027 *Bioengineering and Biotechnology* [Internet] 8. Available from:
1028 https://www.frontiersin.org/articles/10.3389/fbioe.2020.00844

1029 Yang Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and*
1030 *Evolution* 24:1586–1591.

1031 Yang Z, Wong WSW, Nielsen R. 2005. Bayes Empirical Bayes Inference of Amino Acid Sites
1032 Under Positive Selection. *Molecular Biology and Evolution* 22:1107–1118.

1033 Yazdi M, Pahlevanlo A, Jamab MS. 2022. Diversity of Yeasts Microbiota in Iranian Doogh and
1034 Influence of Kluyveromyces Marxianus on its Protein profiles. *Applied Food*
1035 *Biotechnology* 9:41–52.

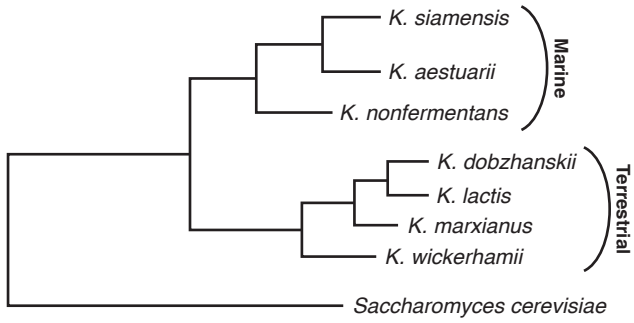
1036 Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an Improved Branch-Site Likelihood Method
1037 for Detecting Positive Selection at the Molecular Level. *Molecular Biology and Evolution*
1038 22:2472–2479.

1039 Zhang L, Zhan H, Xu W, Yan S, Ng SC. 2021. The role of gut mycobiome in health and
1040 diseases. *Therap Adv Gastroenterol* 14:17562848211047130.

1041

Figure 1

A


K. siamensis
K. aestuarii
K. nonfermentans
Marine
K. dobzhanskii
K. lactis
K. marxianus
K. wickerhamii
Terrestrial
Saccharomyces cerevisiae

B

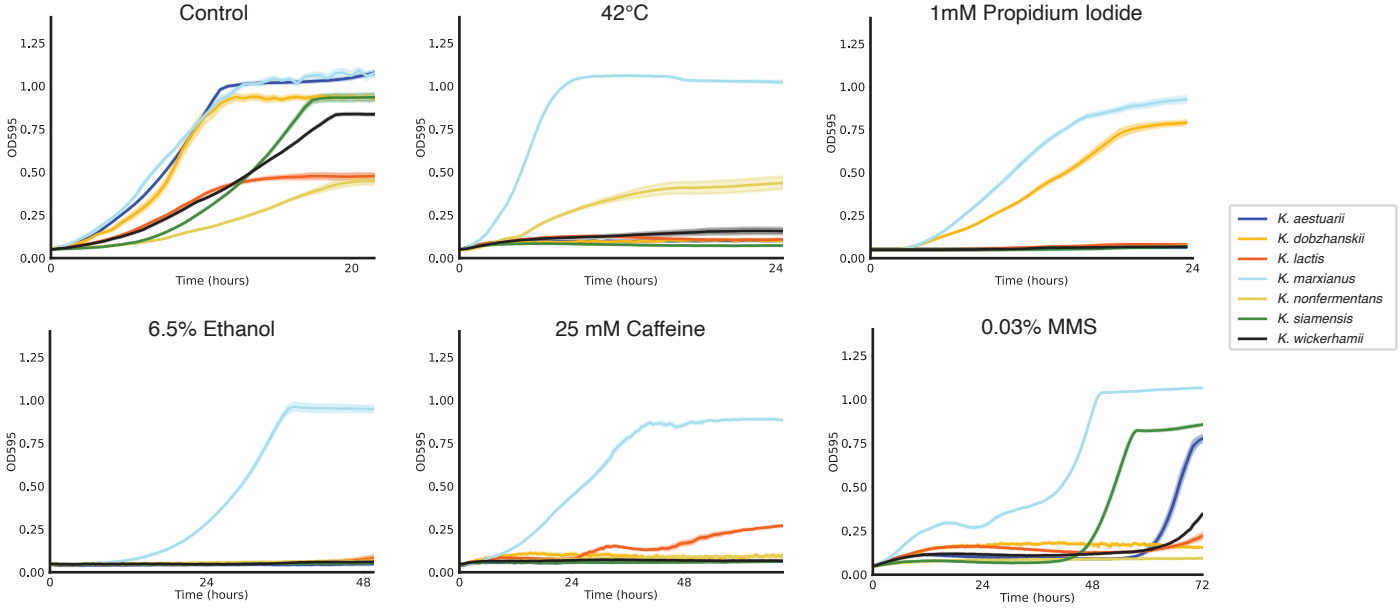
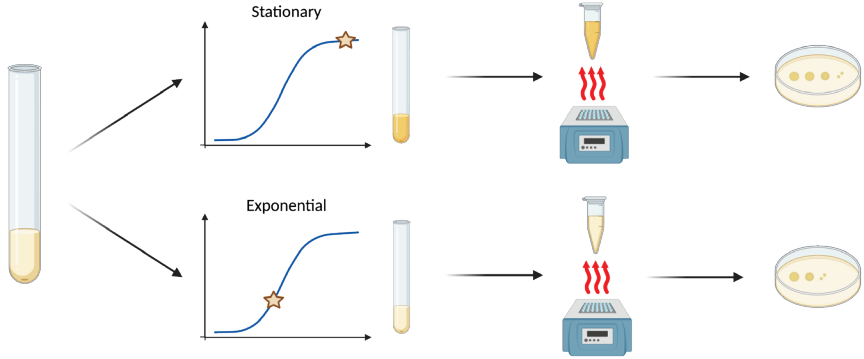
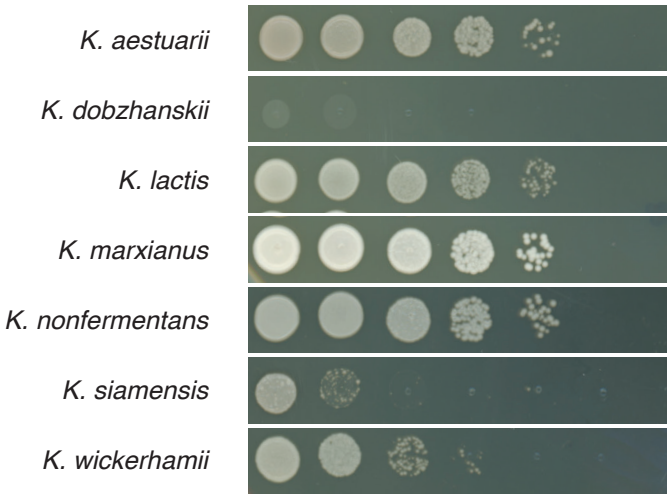

Control
42°C
1mM Propidium Iodide
6.5% Ethanol
25 mM Caffeine
0.03% MMS
OD595
Time (hours)
K. aestuarii
K. dobzhanskii
K. lactis
K. marxianus
K. nonfermentans
K. siamensis
K. wickerhamii

Figure 2

A


Stationary
Exponential

B


Stationary
K. aestuarii
K. dobzhanskii
K. lactis
K. marxianus
K. nonfermentans
K. siamensis
K. wickerhamii

C

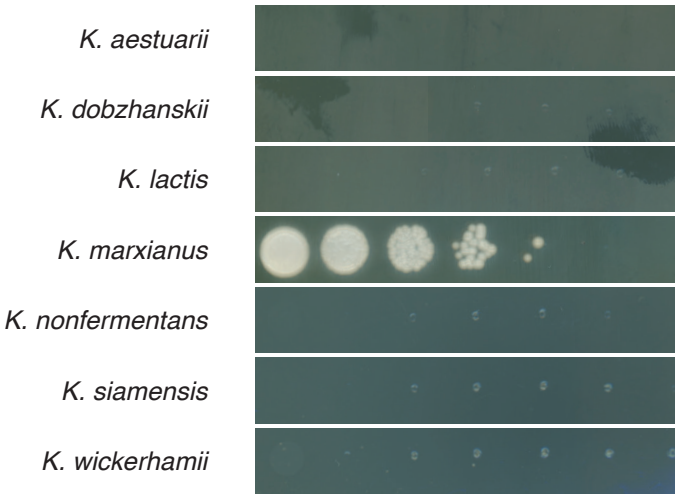

Exponential
K. aestuarii
K. dobzhanskii
K. lactis
K. marxianus
K. nonfermentans
K. siamensis
K. wickerhamii

Figure 3

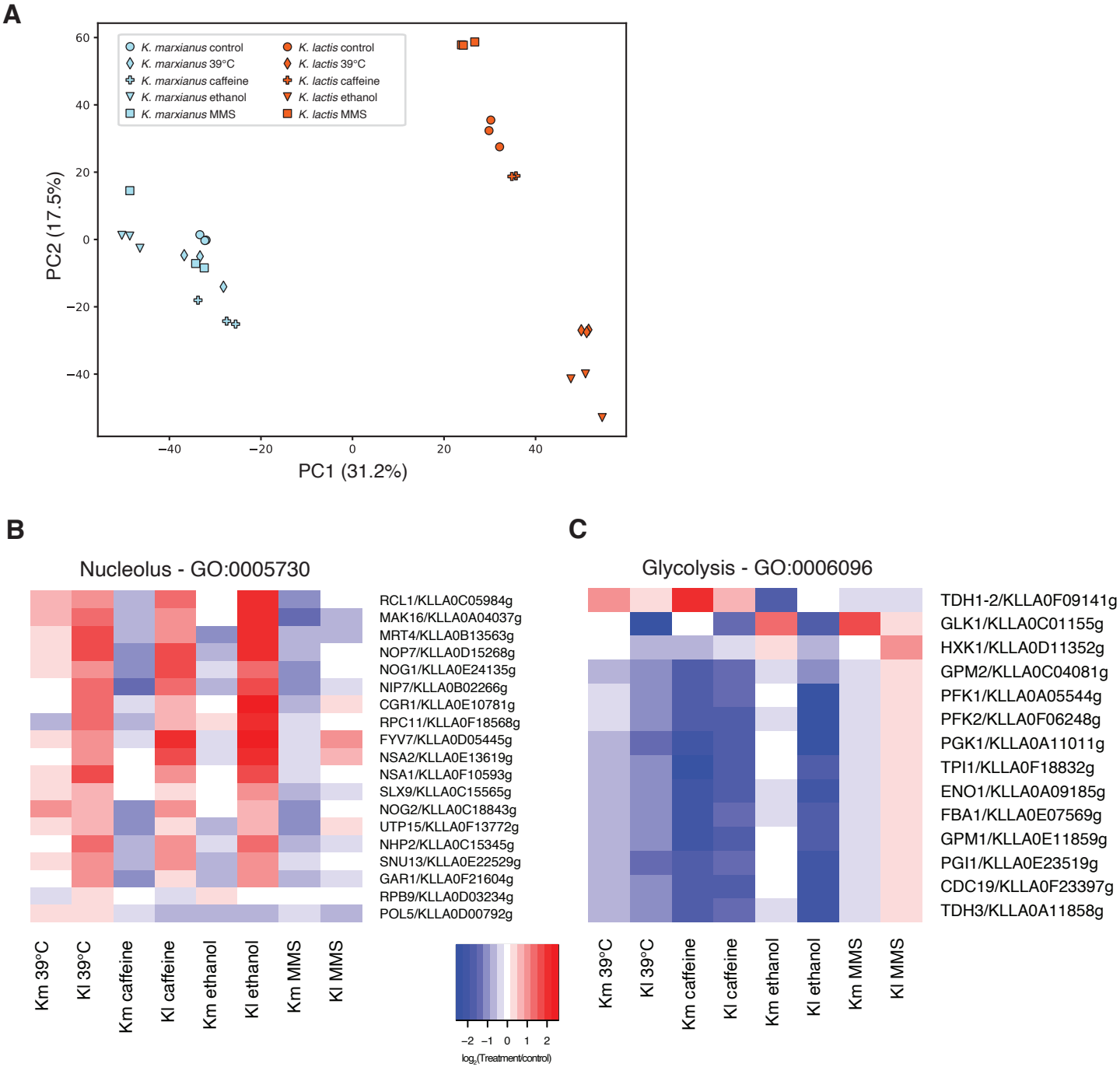
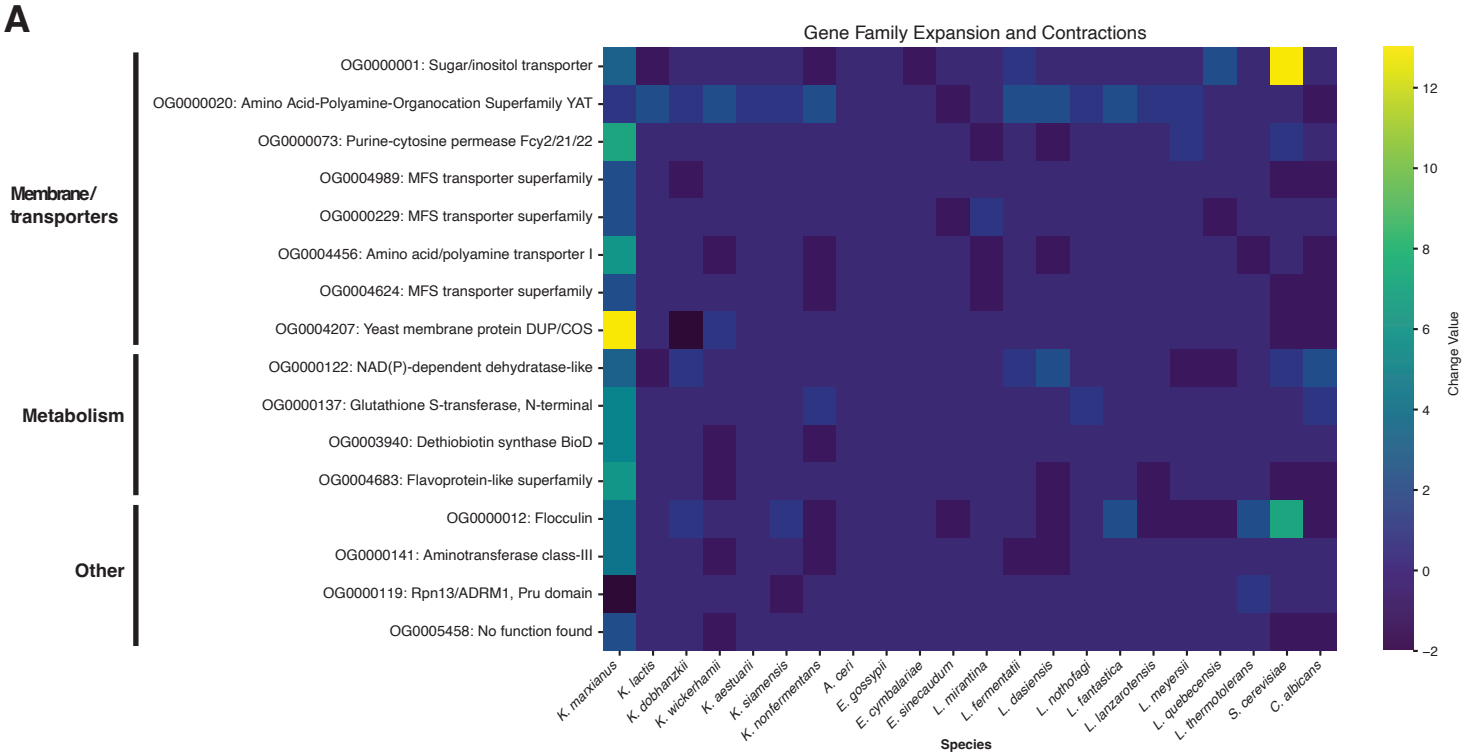

A
K. marxianus control
K. marxianus 39°C
K. marxianus caffeine
K. marxianus ethanol
K. marxianus MMS
K. lactis control
K. lactis 39°C
K. lactis caffeine
K. lactis ethanol
K. lactis MMS
PC2 (17.5%)
PC1 (31.2%)
B
Nucleolus - GO:0005730
RCL1/KLLA0C05984g
MAK16/KLLA0A04037g
MRT4/KLLA0B13563g
NOP7/KLLA0D15268g
NOG1/KLLA0E24135g
NIP7/KLLA0B02266g
CGR1/KLLA0E10781g
RPC11/KLLA0F18568g
FYV7/KLLA0D05445g
NSA2/KLLA0E13619g
NSA1/KLLA0F10593g
SLX9/KLLA0C15565g
NOG2/KLLA0C18843g
UTP15/KLLA0F13772g
NHP2/KLLA0C15345g
SNU13/KLLA0E22529g
GAR1/KLLA0F21604g
RPB9/KLLA0D03234g
POL5/KLLA0D00792g
Km 39°C
Kl 39°C
Km caffeine
Kl caffeine
Km ethanol
Kl ethanol
Km MMS
Kl MMS
log2(Treatment/control)
C
Glycolysis - GO:0006096
TDH1-2/KLLA0F09141g
GLK1/KLLA0C01155g
HXK1/KLLA0D11352g
GPM2/KLLA0C04081g
PFK1/KLLA0A05544g
PFK2/KLLA0F06248g
PGK1/KLLA0A11011g
TPI1/KLLA0F18832g
ENO1/KLLA0A09185g
FBA1/KLLA0E07569g
GPM1/KLLA0E11859g
PGI1/KLLA0E23519g
CDC19/KLLA0F23397g
TDH3/KLLA0A11858g
Km 39°C
Kl 39°C
Km caffeine
Kl caffeine
Km ethanol
Kl ethanol
Km MMS
Kl MMS

Figure 4


A
Gene Family Expansion and Contractions
Membrane/ transporters
Metabolism
Other
OG0000001: Sugar/inositol transporter
OG0000020: Amino Acid-Polyamine-Organocation Superfamily YAT
OG0000073: Purine-cytosine permease Fcy2/21/22
OG0004989: MFS transporter superfamily
OG0000229: MFS transporter superfamily
OG0004456: Amino acid/polyamine transporter I
OG0004624: MFS transporter superfamily
OG0004207: Yeast membrane protein DUP/COS
OG0000122: NAD(P)-dependent dehydratase-like
OG0000137: Glutathione S-transferase, N-terminal
OG0003940: Dethiobiotin synthase BioD
OG0004683: Flavoprotein-like superfamily
OG0000012: Flocculin
OG0000141: Aminotransferase class-III
OG0000119: Rpn13/ADRM1, Pru domain
OG0005458: No function found
K. marxianus
K. lactis
K. dobhanzkii
K. wickerhamii
K. aestuarii
K. siamensis
K. nonfermentans
A. ceri
E. gossypii
E. cymbalariae
E. sinecaudum
L. mirantina
L. fermentatii
L. dasiensis
L. nothofagi
L. fantastica
L. lanzarotensis
L. meyersii
L. quebecensis
L. thermotolerans
S. cerevisiae
C. albicans
Species
Change Value
12
10
8
6
4
2
0
−2

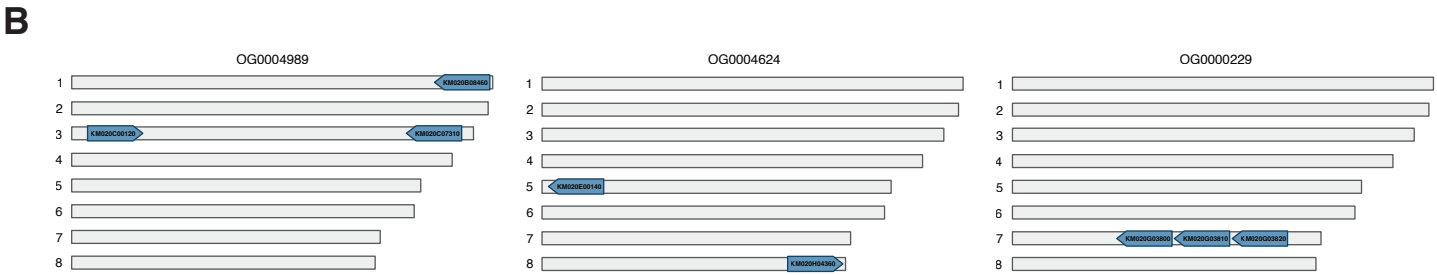
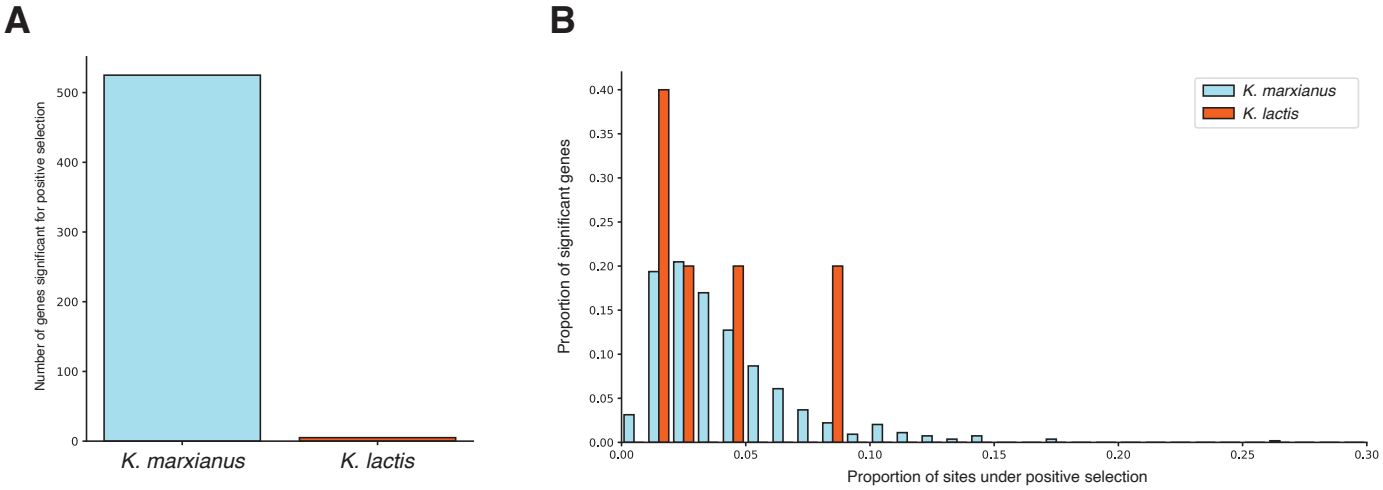
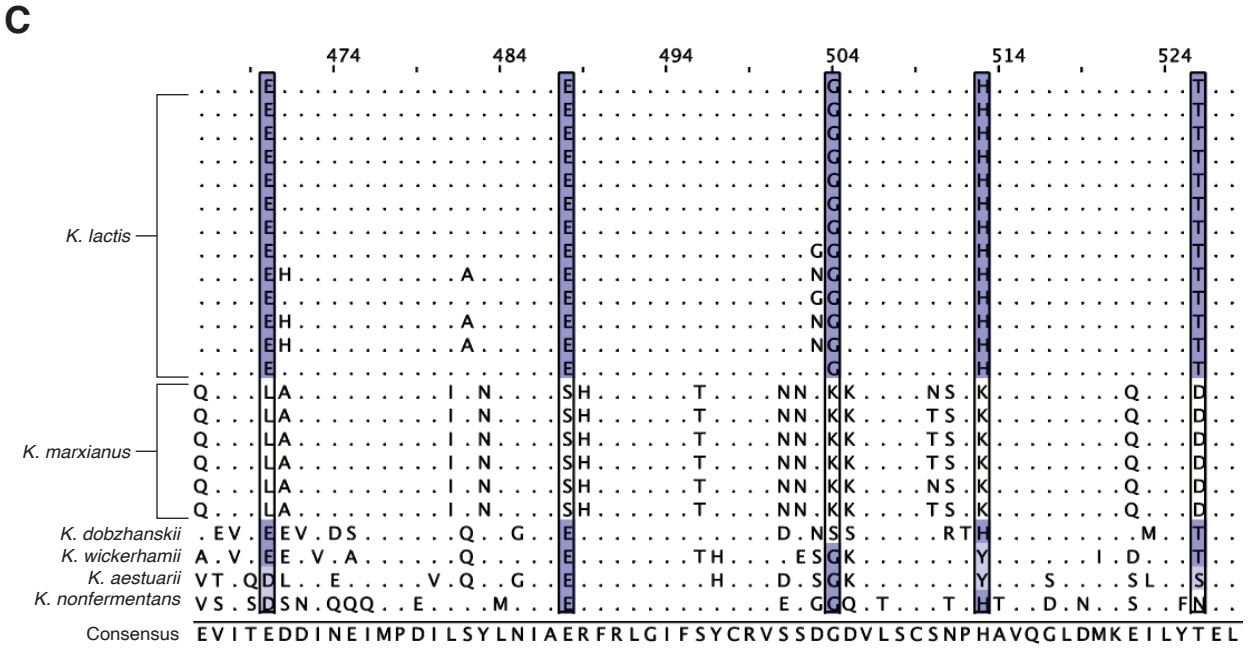

B
OG0004989
OG0004624
OG0000229
1
2
3
4
5
6
7
8
KM020B08460
KM020C00120
KM020C07310
KM020E00140
KM020H04360
KM020G03800
KM020G03810
KM020G03820

Figure 5





D

K. lactis, K. marxianus genes	NI	*Saccharomyces* gene	Standard Name	Function
KLLA0E09329g, KM020H00950	0.19	YOR020W-A	MCO10	Subunit I of the mitochondrial F1F0 ATP synthase
KLLA0E07569g, KM020D04120	0.26	YKL060C	FBA1	Fructose 1,6-bisphosphate aldolase
KLLA0D10780g, KM020E01340	0.32	YJR118C	ILM1	Unknown function, possible mitochondrial involvement
KLLA0C12133g, KM020H02920	0.39	YNL284C	MRPL10	Mitochondrial ribosomal protein
KLLA0A10175g, KM020A03410	0.47	YMR313C	TGL3	Triacylglycerol lipase and LPE acyltransferase
KLLA0F19162g, KM020D01640	0.49	YDR111C	ALT2	Catalytically inactive alanine transaminase
KLLA0B04994g, KM020F01720	0.50	YLR443W	ECM7	Putative integral membrane protein, calcium uptake
KLLA0F00528g, KM020F00240	0.51	YAL061W	BDH2	Putative medium-chain alcohol dehydrogenase
KLLA0B02607g, KM020D05580	0.59	YKL187C	FAT3	Fatty acid uptake
KLLA0D16236g, KM020F04880	0.63	YGR130C	N/A	Component of the eisosome
KLLA0F26048g, KM020E04040	0.68	YMR066W	SOV1	Mitochondrial protein of unknown function

E

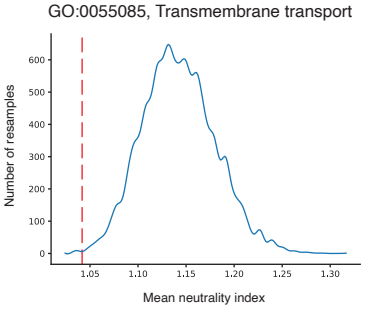


Figure S1

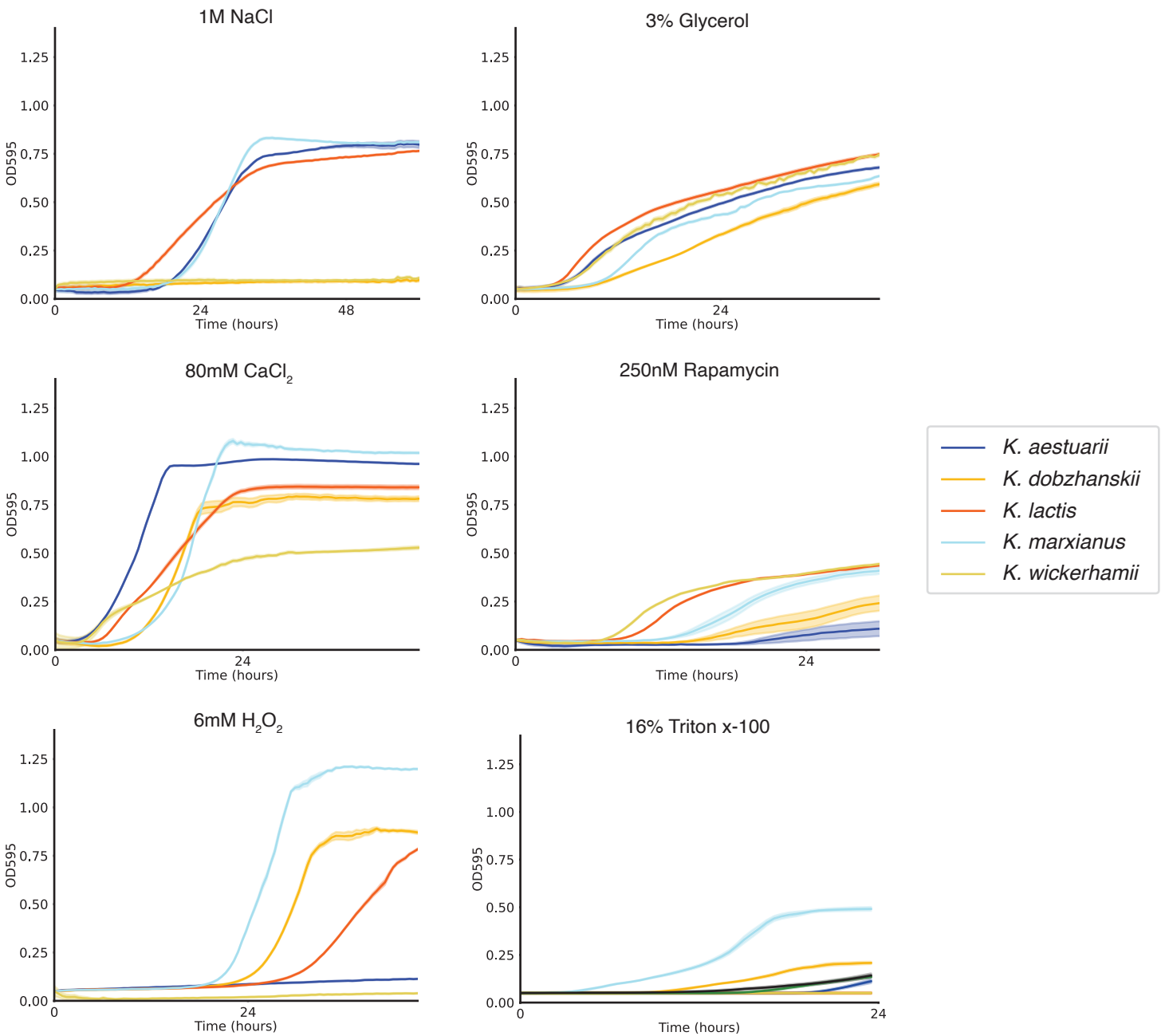

1M NaCl
3% Glycerol
80mM $CaCl_2$
250nM Rapamycin
6mM H_2O_2
16% Triton x-100
OD595
Time (hours)
0.00
0.25
0.50
0.75
1.00
1.25
0
24
48
K. aestuarii
K. dobzhanskii
K. lactis
K. marxianus
K. wickerhamii

Figure S2

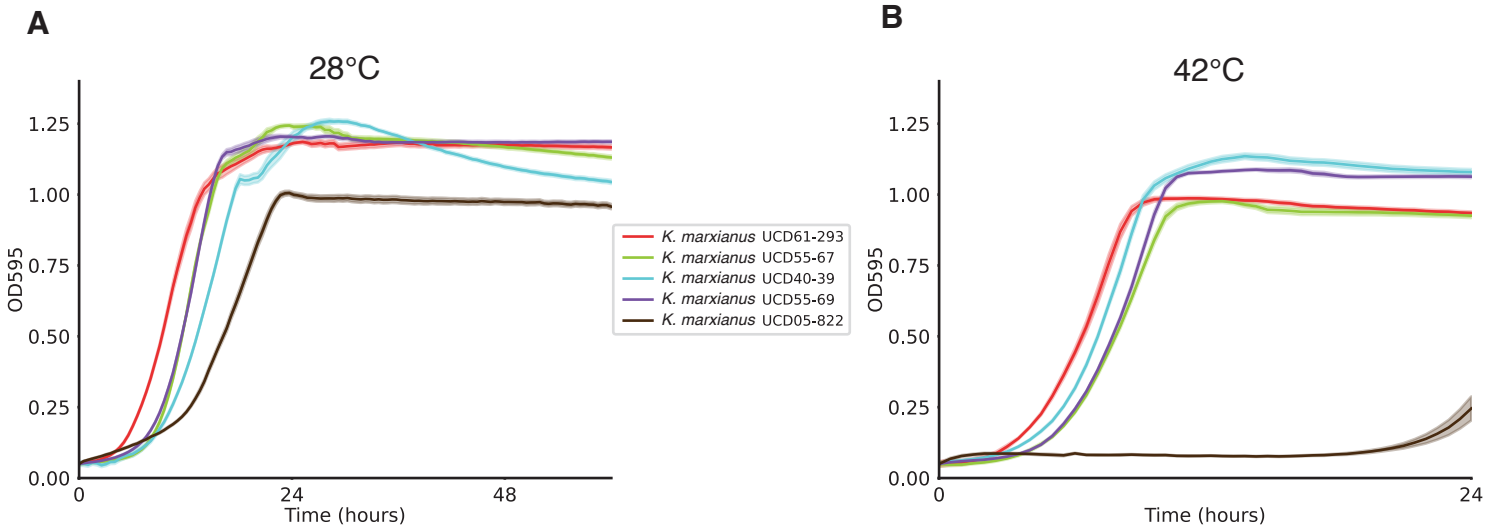

A
28°C
OD595
1.25
1.00
0.75
0.50
0.25
0.00
0
24
48
Time (hours)
K. marxianus UCD61-293
K. marxianus UCD55-67
K. marxianus UCD40-39
K. marxianus UCD55-69
K. marxianus UCD05-822
B
42°C
OD595
1.25
1.00
0.75
0.50
0.25
0.00
0
24
Time (hours)

Figure S3

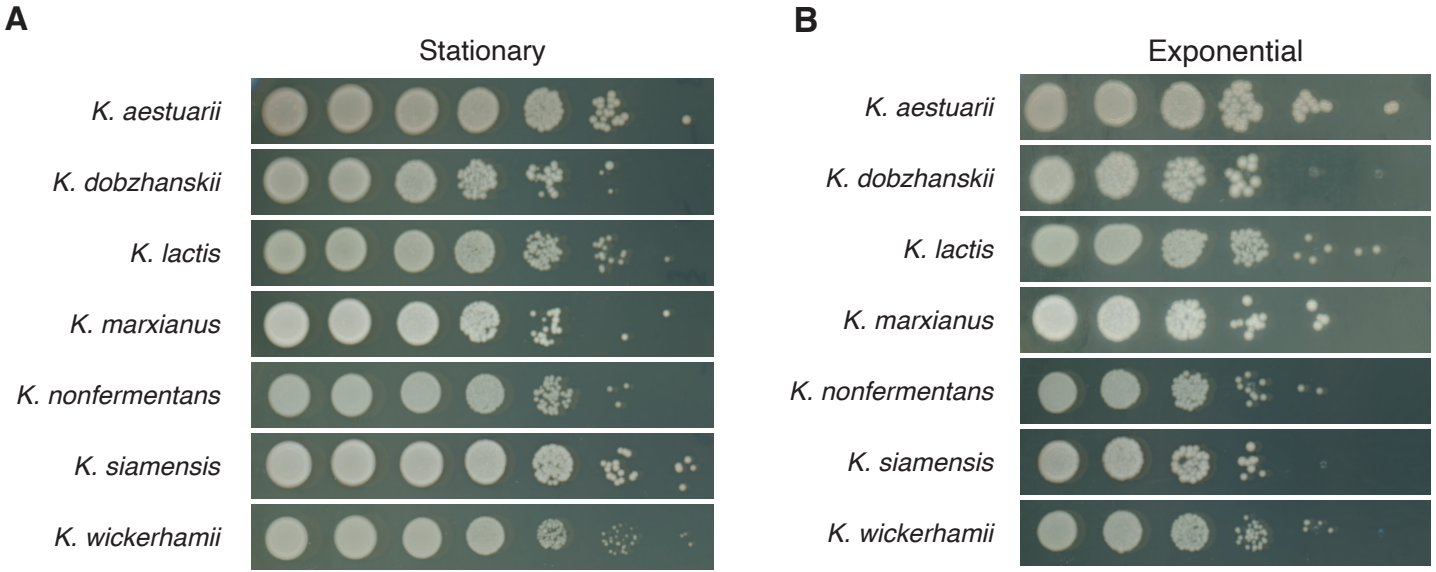

A
Stationary
K. aestuarii
K. dobzhanskii
K. lactis
K. marxianus
K. nonfermentans
K. siamensis
K. wickerhamii
B
Exponential
K. aestuarii
K. dobzhanskii
K. lactis
K. marxianus
K. nonfermentans
K. siamensis
K. wickerhamii

Figure S4

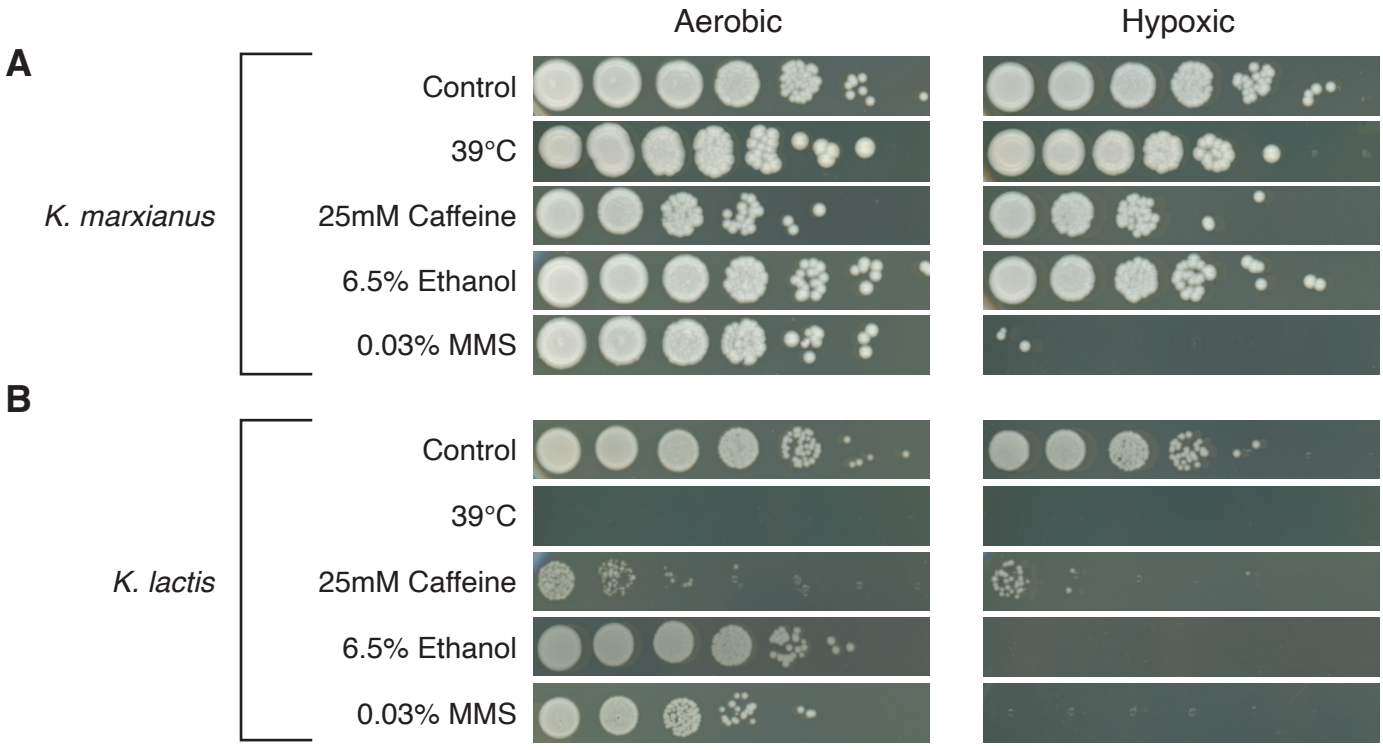
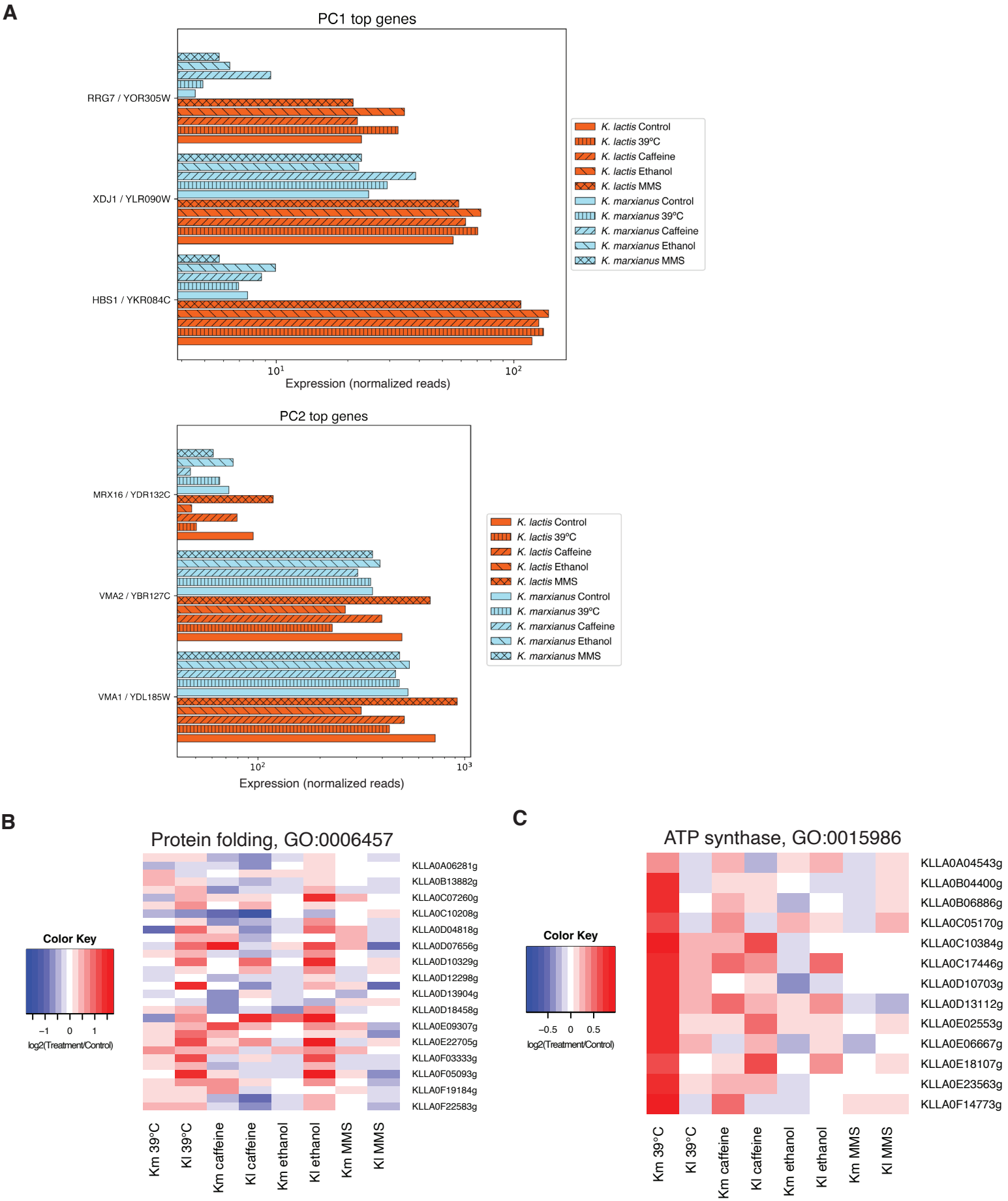

Aerobic
Hypoxic
A
K. marxianus
Control
39°C
25mM Caffeine
6.5% Ethanol
0.03% MMS
B
K. lactis
Control
39°C
25mM Caffeine
6.5% Ethanol
0.03% MMS

Figure S5


A
PC1 top genes
RRG7 / YOR305W
XDJ1 / YLR090W
HBS1 / YKR084C
10^1
10^2
Expression (normalized reads)
K. lactis Control
K. lactis 39°C
K. lactis Caffeine
K. lactis Ethanol
K. lactis MMS
K. marxianus Control
K. marxianus 39°C
K. marxianus Caffeine
K. marxianus Ethanol
K. marxianus MMS
PC2 top genes
MRX16 / YDR132C
VMA2 / YBR127C
VMA1 / YDL185W
10^2
10^3
Expression (normalized reads)
B
Protein folding, GO:0006457
Color Key
−1 0 1
log2(Treatment/Control)
KLLA0A06281g
KLLA0B13882g
KLLA0C07260g
KLLA0C10208g
KLLA0D04818g
KLLA0D07656g
KLLA0D10329g
KLLA0D12298g
KLLA0D13904g
KLLA0D18458g
KLLA0E09307g
KLLA0E22705g
KLLA0F03333g
KLLA0F05093g
KLLA0F19184g
KLLA0F22583g
Km 39°C
Kl 39°C
Km caffeine
Kl caffeine
Km ethanol
Kl ethanol
Km MMS
Kl MMS
C
ATP synthase, GO:0015986
Color Key
−0.5 0 0.5
log2(Treatment/Control)
KLLA0A04543g
KLLA0B04400g
KLLA0B06886g
KLLA0C05170g
KLLA0C10384g
KLLA0C17446g
KLLA0D10703g
KLLA0D13112g
KLLA0E02553g
KLLA0E06667g
KLLA0E18107g
KLLA0E23563g
KLLA0F14773g
Km 39°C
Kl 39°C
Km caffeine
Kl caffeine
Km ethanol
Kl ethanol
Km MMS
Kl MMS