Rapid, scalable, combinatorial genome engineering by Marker-less Enrichment and Recombination of Genetically Engineered loci (MERGE)

Authors:

Mudabir Abdullah[1], Brittany M. Greco[1], Jon M. Laurent[2†], Michelle Vandeloo[1], Edward M. Marcotte[3], Aashiq H. Kachroo[1,‡]

Affiliations:

[1]Centre for Applied Synthetic Biology, Department of Biology, 7141 Sherbrooke St. W, Concordia University, Montreal, QC, Canada

[2]Institute of Systems Genetics, NYU Langone Health, New York, USA

[3]Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX, USA

[†]Current address: Pandemic Response Lab, 45-18 Ct Square West, Long Island City, NY 11101

[‡]Correspondence to *aashiq.kachroo@concordia.ca*

Keywords: Genome engineering, CRISPR-Cas9, humanized yeast, proteasome

Abbreviations: CFU, Colony Forming Units; DSB, Double-Strand Break; HDR, Homology-directed DNA Repair; HR, Homologous Recombination; CELECT, CRISPR-Cas9-based selection to enrich genotypes; MERGE, Marker-less Enrichment and Recombination of Genetically Engineered loci; SGA, Synthetic Genetic Array

Abstract:
Large-scale genome engineering in yeast is feasible primarily due to prodigious homology-directed DNA repair (HDR), a plethora of genetic tools, and simple conversion between haploid and diploid forms. However, a major challenge to rationally building multi-gene processes in yeast arises due to the combinatorics of combining all of the individual edits into the same strain. Here, we present an approach for scalable, precise, multi-site genome editing that combines all edits into a single strain without the need for selection markers by using CRISPR-Cas9 and gene drives. First, we show that engineered loci become resistant to the corresponding CRISPR reagent, allowing the enrichment of distinct genotypes. Next, we demonstrate a highly efficient gene drive that selectively eliminates specific loci by integrating CRISPR-Cas9 mediated Double-Strand Break (DSB) generation and homology-directed recombination with yeast sexual assortment. The method enables **M**arker-less **E**nrichment and **R**ecombination of **G**enetically **E**ngineered loci (MERGE) in yeast. We show that MERGE converts single heterologous yeast loci to homozygous loci at ~100% efficiency, independent of chromosomal location. Furthermore, MERGE is equally efficient at converting and combining loci, thus identifying viable intermediate genotypes. Finally, we establish the feasibility of MERGE by engineering a fungal carotenoid biosynthesis pathway and most of the human α proteasome core into yeast. MERGE, therefore, lays the foundation for marker-less, highly efficient, and scalable combinatorial genome editing in yeast.

Introduction:

Baker's yeast has long served as a convenient chassis for bioengineering owing to its genetic tractability, versatile metabolism, and ease of culture in the lab. Decades of fundamental research, together with the development of high throughput toolkits and genome engineering capacities, have established yeast as an ideal model eukaryote for system genetics and synthetic biology[1]. The availability of many selectable genetic markers and simple conversion between haploid and diploid forms has provided avenues to easily combine pairs of genetically engineered loci to understand gene-gene interactions at a global scale[2]. For more extensive genetic alterations, yeast's highly efficient Homologous Recombination (HR) pathway even enables the synthesis of entire chromosomes, although this approach requires iterative use of selection markers and tedious repetitive procedures[3]. Nonetheless, the ability to alter large contiguous segments of genomic loci has many applications, such as genome minimization[4], multiplex genome editing[5], and the total synthesis of the *Mycoplasma* and *E. coli* genomes using yeast HR[6,7].

In spite of the progress in whole genome engineering, editing intermediate numbers of independent genomic loci–greater than two and in discontiguous regions of the genome–still presents a significant challenge. While strategies exist for *E. coli*[8] (e.g., MAGE), diploid organisms such as yeast present the additional editing challenge of multiple alleles for each (independently assorting) genomic locus. Moreover, while high throughput cloning strategies and the reduced cost of *de novo* DNA synthesis now allow swapping entire heterologous pathways or protein complexes into yeast, such efforts frequently require the deletion of corresponding yeast loci, such as the case for efforts to systematically humanize yeast genes[9,10], which may entail replacing genes at their corresponding genomic loci to retain native regulation. Beyond these aspects, the expression of each new gene often reveals incompatibilities associated with the engineered pathway. Thus, there is a need for rapid, multi-site, progressive genome editing strategies to address these issues.

The efficiency and speed of CRISPR-Cas9-based genome engineering allow straightforward editing of multiple yeast loci in a single strain, eliminating the need for markers[11–15]. However, the approach gets progressively more challenging for multi-gene systems when the fitness of the intermediate genotypes is unknown (**Figure S1**). Therefore, rationally building heterologous genetic modules in a yeast surrogate requires a highly scalable and combinatorial genome editing technology.

Synthetic Genetic Array (SGA) analysis permits the combination of loci in a fitness-driven manner, requiring markers linked to the modified loci and haploid-specific selections[16]. Using an SGA-like strategy to build a heterologous multi-gene system would need several unique markers linked to each gene. However, the lack of adequate selection markers limits its application. The Green Monster (GM) method bypasses the marker dependency by using Green Fluorescent Protein (GFP) expression as a readout to combine engineered loci[17]. However, implementing the GM strategy can be challenging for heterologous pathway engineering with an added burden of expressing fluorescent cassettes[18]. Thus, accomplishing large-scale combinatorial genome editing in yeast necessitates new genetic tools that circumvent these requirements.

In this work, we describe a novel, CRISPR-Cas9-based method to readily combine genetically engineered loci without the need for markers and exogenous repair templates. The approach involves generating CRISPR-Cas9-mediated Double-Strand Breaks (DSBs) and highly

efficient Homology Directed Repair (HDR) along with yeast mating and sporulation to randomly assort edited sites performing multi-site and marker-less combinations of engineered loci. This selective and successive elimination of specific yeast loci mimics a gene drive[19–22] (**Figure S2**) and facilitates Marker-less Enrichment and Recombination of Genetically Engineered yeast loci (**MERGE**).

We show that CRISPR-mediated selection operates similarly to classical selection markers, enabling MERGE to efficiently explore many combinations of genetically engineered loci, revealing a fitness-driven path to engineering any heterologous system in yeast. We further demonstrate that MERGE enables rapid assembly of an entire carotenoid biosynthesis pathway by performing a multi-site and marker-less combination of distinct engineered loci. Finally, using a multiplexed version of MERGE, we humanize a near-complete α proteasome core (6 of 7 subunits) in yeast while revealing a fitness-driven path to the humanization of complex processes.

Results & Discussion:

CRISPR-Cas9 allows marker-less selection and enrichment of unique genotypes in yeast (CELECT)

CRISPR-Cas9 enables the precise and marker-less editing of both essential and non-essential yeast loci owing to significantly lower error-prone Non-Homologous End Joining (NHEJ) in yeast relative to HDR[23]. Therefore, Cas9-sgRNA-induced lethality serves as a rapid test for a functional CRISPR reagent in yeast (Colony Forming Units Observed or $\mathbf{CFU_O}$ = ~0) compared to the Cas9 alone (Colony Forming Units Expected or $\mathbf{CFU_E}$). To verify the ON-target activity of CRISPR reagents, the transformation of a plasmid carrying expression cassettes for both Cas9 and a sgRNA targeting a specific locus (**pCas9-sgRNA^{locus}**) in a yeast strain that harbors a corresponding engineered locus resistant to further targeting allows the survival of colonies similar to the pCas9 alone ($\mathbf{CFU_O/CFU_E}$ **= ~1**) (**Figure 1A**).

We tested this strategy by humanizing α proteasome core genes in yeast that are functionally replaceable by their human counterparts[24]. Using CRISPR-Cas9, we replaced each yeast α proteasome gene with its human ortholog at the native loci (**Figure S3**). Additionally, we modified non-essential loci to test if the CRISPR-mediated selection is broadly applicable and found that all engineered strains are resistant to the corresponding CRISPR-Cas9-sgRNA mediated lethality ($\mathbf{CFU_O/CFU_E}$ **= ~1, Figure 1B, Figure S4A & S4B**). Henceforth, we refer to the CRISPR-Cas9-mediated selection of genotypes as CELECT.

To further verify if the resistance of humanized strains is exclusive to its corresponding CRISPR reagent, all single-humanized α proteasome strains were mixed in culture (**Figure 1C**). To quantify the enrichment of unique genotypes, we inoculated a *ade2Δ::kanMX* haploid strain in the mixture. The **pCas9-sgRNA^{ADE2}** exclusively enriched for resistant *ade2Δ* genotype, whereas all other genotypes in the mix harboring a wild type *ADE2* locus are inviable (**Figure 1C-ii**). Conversely, the transformation of **pCas9-sgRNA^{RPT5}** targeting *RPT5* (a base subunit of the proteasome complex), for which all the strains in the mix harbor a wild-type copy, shows no survivors (**Figure 1C-iii**). We demonstrate that each CRISPR reagent targeting yeast α proteasome genes selected a corresponding humanized genotype from a mix, respectively

(**Figure 1C-iv**). Thus, CRISPR-Cas9-mediated resistance functions similarly to conventional antibiotic or auxotrophic markers in yeast.

MERGE^0 is nearly 100% efficient at converting loci irrespective of the yeast gene location on the chromosome

Mating of engineered haploid strains with wild-type generates a heterozygous genotype. CRISPR-Cas9 reagent targeted to either allele should enable the conversion to homozygous diploid at high efficiency while also enriching the desired genotype (**MERGE level 0** or **MERGE^0**) (**Figure 2A**). We used *ADE2/ade2Δ::kanMX* heterozygous knockout diploid (hetKO) strain to quantify the efficiency of **MERGE^0**. The wild type *ADE2* allele is susceptible to **pCas9-sgRNA^{ADE2}** mediated DSB. In contrast, the *ade2Δ::kanMX* allele is resistant, proving a readout of conversion as the loss-of-function of *ADE2* results in a red color colony phenotype.

The transformation of the wild-type haploid or diploid yeast with **pCas9-sgRNA^{ADE2}** showed a lethal phenotype (~0-20 **CFU$_O$** on average) (**Figure 2B**). However, **pCas9-sgRNA^{ADE2}** transformation in *ADE2*/*ade2Δ*::*kanMX* hetKO strain shows resistance while also converting the locus to *Δade2*::*KanMX* at ~100% efficiency (**Figures 2B & S5A**). **MERGE^0** performs with a comparable proficiency at single-humanized α proteasome and landing pad loci (**Figures 2C, S5B & S5C**). To further test if **MERGE^0** can convert any heterozygous yeast locus independent of the position on a chromosome, we explored the strategy across many yeast genes located on chromosome 1. We designed a CRISPR reagent to target *KanMX* (**pCAS9-sgRNA^{KanMX}**) in hetKO diploid strains harboring the *KanMX* cassette instead of a yeast gene. **MERGE^0** converted all heterozygous loci to homozygous wild-type alleles, respectively, with simultaneous loss of *KanMX* cassette (except for *CNE1*) (**Figure 2D, S6A & S6B**).

Furthermore, using CRISPR-Cas9 to target a wild-type yeast locus instead of the *KanMX* cassette in a hetKO diploid strains, **MERGE^0** enables single-step gene essentiality screening in yeast, significantly increasing efficiency compared to the classical methods such as Tetrad dissection or SGA analysis[25,26]. Except for a non-essential *α3* gene, all six essential α proteasome genes are inviable post-transformation of the corresponding CRISPR reagents (**Figure 2E**).

MERGE^1 permits a fitness-driven combination of the engineered loci

Given the high efficiency of **MERGE^0** at converting one yeast locus, we next tested if the method works similarly to convert and combine two distinct loci (**MERGE level 1 or MERGE^1**). Mating yeast strains, each with one modified locus, facilitate the combination as heterozygotes at two separate loci. CRISPR-Cas9 with two sgRNAs that target each unique locus allows simultaneous conversion to paired CRISPR-resistant loci, respectively (**Figure 3A**). To test the pipeline across all humanized *α* proteasome strains, we first used **MERGE^0** to move all singly humanized *α* proteasome loci from BY4741 background to SGA strain background in both mating-types (**Figure S4B**). To verify if **MERGE^1** can simultaneously convert two distinct loci, we tested the efficiency of combining paired-humanized α3/α7 and α1/α7 genotypes. We show that **MERGE^1** is highly efficient as each heterozygous genotype is resistant to the double-sgRNA CRISPR-Cas9 reagent while converting and combining two humanized loci (**Figure 3B, S7A**). The CRISPR reagent is lethal in wild-type and singly humanized strains (**Figure 3B & S7B**). The sporulation of viable diploids yields haploid double-humanized strains in both mating-types. Genotyping

148 randomly picked colonies confirms the conversion to humanized loci (**Figure 3B**). Additionally,
149 the double-humanized α7/α1 genotype, while viable, shows a sporulation defect (a phenotype
150 associated with *Hsα7*) (**Figure S7A**). Thus, **MERGE^{1}** allows the survival and enrichment of only
151 viable homozygous double-humanized loci without the requirement of any diploid-specific
152 selection. However, a double-humanized α5/α7 is inviable as a combined genotype, as evidenced
153 by double-sgRNA CRISPR-Cas9-mediated lethality (**CFU_O/CFU_E = ~0**). PCR-based genotyping
154 of a surviving colony shows homozygosity of only the Hsα7 locus (**Figure 3C**). In contrast, while
155 a single-humanized Hs*α6* strain is temperature-sensitive (TS) at 37°C, combination with the
156 neighboring humanized α7 gene rescues the phenotype (**Figure S7A′**). However, the TS
157 phenotype is associated with one of the variants of Hsα6 (Variant 1, 37>Glycine) (**Figure S8A**),
158 whereas another common variant (Variant 2, 37>Valine) shows no growth defect at 37°C (**Figure**
159 **S8B**). Thus, **MERGE^{1}** revealed the fitness of paired-humanized genotypes similar to synthetic
160 genetic interactions without the need for linked markers or haploid-specific selections[16].
161 To systematically determine if there are specific pairwise restrictions to the humanization
162 of α-subunits, we mated all haploid single-humanized strains obtaining diploid heterozygous
163 genotypes (21 different genotypes). The corresponding CRISPR-Cas9-based selection and a
164 simple readout of **CFU_O/CFU_E** identified the permitted double-humanized genotypes (17/21
165 genotypes). The data reveals that only specific double-humanized genotypes are viable, whereas
166 some are not (**Figure 3D & S7C**). The incompatibility of paired genotypes comprising *Hsα1/α6*,
167 *Hsα4/α6*, *Hsα5/α6*, and *Hsα5/α7* may likely be due to the missing neighboring interactions within
168 the α proteasome core, except in the case of *Hsα5,α6* pair (**Figure S7C′**). A sequential editing
169 strategy successfully engineered some paired-humanized genotypes to confirm that **MERGE^{1}**
170 represents viable/fit paired genotypes. However, it did not provide a clear perspective of
171 incompatibilities (**Figure S7D**).
172 Given the success of **MERGE^{1}**, we examined the strategy to select double-humanized
173 genotypes randomly from a mixture (Mix, Mate & MERGE or **MERGE$^{M\&M}$**) (**Figure 3E**). We
174 inoculated single-humanized *α1, α3* and *α7* haploid strains of both mating-types as a mix, allowing
175 random mating. Each double-sgRNA CRISPR-Cas9 selection enriched the corresponding paired
176 genotype from the mated mix while converting the wild-type yeast to humanized loci (**Figure 3E-**
177 **iii & 3E-iv**). In mixed culture, the transformation of **pCas9-sgRNA$^{Sc\text{-}\alpha1,\alpha6}$,** a selection for a non-
178 existing genotype, does not yield any viable genotype (**Figure 3F-ii**). Therefore, **MERGE$^{M\&M}$** can
179 be scaled to obtain several paired genotype combinations of engineered loci from a mix.
180

181 **MERGE is scalable to combine multiple genetically engineered loci**

182 The CRISPR-Cas9 system enables multiplexed editing by introducing multiple sgRNAs[27].
183 To test scalability, we designed **MERGE^{MX}** (**MERGE, mate and multiplex**) to verify if >2 genetic
184 loci can simultaneously convert to engineered loci by building the 4-gene carotenoid biosynthesis
185 pathway from the carotenogenic yeast *Xanthophyllomyces dendrorhous* into Baker's yeast
186 (**Figure 4A**). The carotenoid pathway provides a colony color readout as a proxy for pathway
187 engineering (**Figure 4A**). **MERGE^{0}** generated haploid strains of opposite mating-types for each
188 carotenoid transcription unit (**Figure S5C**), **MERGE^{1}** provided the double-carotenoid genotypes
189 (**Figure 4B**), and **MERGE^{MX}** enriched a complete homozygous carotenoid pathway genotype
190 (**Figure 4C**). Furthermore, genotyping of randomly picked dark orange colonies confirmed the
191 conversion and combination of the engineered loci (**Figure 4C**).

To test whether **MERGE^{MX}** can perform the combination of >2 humanized *α* proteasome genotypes, we mated a previously obtained Hs*α1α2α3* strain (**MERGE^{1}**) with a wild-type type strain generating a heterozygous diploid for all three loci. The strain provided a platform to test **MERGE^{MX}** by simultaneously combining 3 distinct humanized loci. The triple-sgRNA CRISPR reagent (**pCas9-sgRNA$^{Sc\text{-}\alpha1,\alpha2,\alpha3}$**) converted all yeast loci to human versions (7 of 9 colonies tested) (**Figure 4D**). Next, using **MERGE^{MX}**, we tested if a triple-humanized Hs*α4α6α7* genotype is viable. Transformation of triple-sgRNA, **pCas9-sgRNA$^{Sc\text{-}\alpha4,\alpha6,\alpha7}$** in a mated mix of haploid Hs*α4α7* (Mat A) and Hs*α6* (Mat *α*) strains allowed simultaneous conversion of three wild-type yeast loci to humanized versions without using any diploid specific selection (**Figure 4E & S11A**). Thus, if a triple-humanized intermediate genotype is viable, **MERGE^{MX}** can enrich and combine the specific genotype.

To further address throughput, we designed a CRISPR reagent targeting a GFP expression cassette in a GM strain[17] (16 GFP loci, **Figure S9A)**. The transformation of **pCas9-sgRNA^{GFP}** resulted in a few survivors that failed to show GFP expression, likely due to mutations at GFP loci due to NHEJ (**Figure S9B & S9C**), suggesting successful targeting of the majority of GFP cassettes.

MERGE enables fitness-driven engineering of a near entire human α proteasome core in yeast

To explore the proficiency of **MERGE** for testing many combinations of engineered loci, we asked if an entire yeast heptameric *α* proteasome ring is humanizable. As an alternate strategy, we also tested the feasibility of sequential engineering using repetitive CRISPR selections and exogenous human gene repair templates (**Figure 5A**). The co-transformation of a triple-sgRNA CRISPR reagent (**pCas9-sgRNA$^{Sc\text{-}\alpha1,\alpha2,\alpha3}$**) targeting yeast *α*1, *α*2 and *α*3 genes and PCR fragments of human gene repair templates was successful in obtaining a triple-humanized strain (**Figure S10A-i**). A similar strategy generated triple-humanized Hs*α1α2α4* and Hs*α1α3α4* strains but failed to obtain a quadruple-humanized Hs*α4α5α6α7* (**Figure S10A-ii**). Thus, yeast genes are replaceable sequentially either alone or as small-scale simultaneous replacements. Using a triple-humanized genotype (Hs*α1α2α3*), the humanization of yeast *α*4 was successful. Next, we explored Hs*α*5, Hs*α*6 and Hs*α*7 humanizations in parallel; however, we obtained only one quadruple-humanized Hs*α1α2α3α4α7* genotype (**Figure 5A, Figure S10B**). The functional replacement of yeast *α*5 or *α*6 was unsuccessful despite repeated attempts. The plasmid-borne expression Hs*α6* in a quintuple-humanized strain (Hs*α1α2α3α4α7*) resulted in a toxic phenotype (no growth), suggesting that further humanizations are incompatible (**Figure 5A**). Overall, while the sequential strategy was partly successful, it failed to reveal if the inability to humanize yeast α core entirely was due to incompatible genotypes or inefficient genome editing, especially as hybrid human-yeast genotypes show growth defects and reduced transformation efficiencies (**Figure S10C**).

However, **MERGE^{MX}** provided a clear readout of incompatible humanized genotypes, readily generating many combinations of humanized genotypes (**Figure 5B**). We first explored several triple-humanized genotypes, obtaining *Hs*α1*α*2*α*3 & *Hs*α4*α*6*α*7 strains (**Figures 4D & 4E**), while *Hs*α4*α*5*α*6 and *Hs*α4*α*5*α*7 genotypes are inviable (**Figures 5B, S11B & S11C**). By mating yeast strains with distinct humanized genotypes, we next explored many higher-order (>3) combinations, obtaining quadruple-humanized Hs*α*1*α*2*α*3*α*4 and Hs*α*1*α*2*α*3*α*7 genotypes. In

comparison, the quadruple-humanized *Hsα*4*α*5*α*6*α*7 genotype is not feasible (**Figure 5B, Figure S12A, S12B, S12C & S12D)**. Subsequent **MERGE** strategies generated viable quintuple-humanized genotypes, *Hsα*1*α*2*α*3*α*6*α*7 and *Hsα*1*α*2*α*3*α*4*α*6, whereas *Hsα*1*α*2*α*3*α*4*α*5 genotype is inviable (**Figures 5B**, **S12E, S12F, & S12G)**. The following **MERGE^{MX}** assay yielded a viable sextuple-humanized *Hsα*1*α*2*α*3*α*4*α*6*α*7 genotype with a delayed growth phenotype (**Figure S11H**). To conclusively verify if the entire yeast *α* proteasome core is humanizable, we mated partially humanized Hs*α*1*α*2*α*3*α*6*α*7 and Hs*α*4*α*5 strains, allowing the combination of all *α* proteasome genes as heterozygous human-yeast genotypes (**Figure 13A**). **MERGE^{MX}** converted yeast *α*1,*α*2, and *α*3 loci to homozygous human alleles. However, the subsequent conversion of the remaining four yeast *α*4*α*5*α*6*α*7 to human loci using **pCas9-sgRNA$^{Sc\text{-}\alpha4,\alpha5,\alpha6,\alpha7}$** did not yield viable colonies suggesting a fully human *α* proteasome core is incompatible (**Figure S13B**). Overall, **MERGE** successfully tested many combined humanized genotypes. This gradual progression from yeast to humanized *α* proteasome core rescued the viability of specific incompatible double-humanized genotypes, suggesting that these subunits are co-humanizable when neighboring interactions are restored (**Figures 3D & S14**). The data reveals that the proteasome subunits *α5* and *α6* are not co-humanizable in yeast (**Figure 5**).

Proteasome biogenesis is a highly regulated process aided by several assembly chaperones[28–30]. In the case of *α* proteasome core, particularly *α*5 and *α*6, subunits interact with assembly chaperones, enabling ordered assembly[31–33]. Furthermore, the *β* core assembly immediately follows the *α* core, and the incompatible interface may now require human *β* subunits in yeast[28,34]. The heterologous expression of human assembly chaperones or human *β* subunits in a humanized *α* core strain may permit the synthesis of a fully human catalytic core particle in yeast. Given the complex assembly and architecture of the proteasome[35], it is challenging to know if there are a limited number of 'paths' to engineer a fully-humanized 20S core particle, due to a rapidly accumulating number of assays to perform (as in **Figure S1**).

The sporulation failure observed in genomically-replaced strains can lead to dead ends while performing MERGE (associated with all humanized Hs*α*7 genotypes, **Figures S11A, S12A & S12F**). We propose two solutions: One, by allowing the strains with heterozygous engineered loci to sporulate without **MERGE**, followed by using **CELECT** to enrich combined haploid genotypes (**Figure S15A**). Alternatively, a sequential strategy can engineer viable genotypes in a haploid strain (**Figures S12F & S15B**). Furthermore, using CRISPR reagents to generate multiple DSBs at several yeast loci could potentially have OFF-target effects. Therefore, we performed whole-genome sequencing of singly humanized (Hs-*α1*) and quintuple-humanized strains (*Hsα*1*α*2*α*3*α4α*7), ruling out OFF-target DSBs and mutations (**Figure S16**).

Conclusions

In conclusion, MERGE offers highly scalable multi-locus genome engineering in diploid yeast cells by using a high-efficiency CRISPR-based gene-drive-like strategy to overcome the independent assortment of unlinked loci. MERGE has the potential to allow systematic functional genomic analysis in other systems lacking sophisticated tools and drive synthetic and systems biology research from engineering heterologous systems to performing multi-site and genome-wide combinatorial editing in yeast, as we demonstrated by engineering 89 independent sites along Chromosome 1, a complete four gene carotenoid biosynthesis pathway, 16 GFP insertions within the same strain, and a majority of the α proteasome genes. By engineering 6 of 7 human

α proteasome core genes in yeast, our work also demonstrates the remarkable degree of functional conservation in the proteasome complex despite over a billion years of evolutionary divergence, extending from a single gene to nearly an entire module. The data confirm our previous observations that humanization seems to be driven by modules of physically or functionally interacting proteins being similarly replaceable[24]. Further characterization of the incompatibilities should reveal novel orthogonal functions or interactions in diverged species. However, pursuing a combinatorial strategy with MERGE along with a sequential strategy in parallel allows one to inform the other about simultaneous replacements that are likely to work. Humanizing all or multiple members of a protein complex will allow a novel approach to learning human biology, including complex assembly, biogenesis, and variant effects on function, investigations of their contributions to disease, and the possibility of seeking therapies for these diseases in the simplified context of a yeast cell.

Acknowledgements

The authors would like to thank Riddhiman Garge and Dan Boutz (University of Texas at Austin) for many suggestions and critical feedback on gene drives and proteasome biochemistry, Charles Boone (University of Toronto) for sharing the SGA-compatible yeast strains, Frederick Roth (University of Toronto) and Harvard University for sharing Green monster strains, Jef Boeke (New York University) and Grant Brown (University of Toronto) for sharing the carotenoid gene yeast expression plasmids, and Maitreya Dunham (University of Washington) for insightful discussions and feedback on performing MERGE across the entire yeast chromosome.

Funding

This research was funded by grants from the Natural Sciences and Engineering Research Council (NSERC) of Canada (Discovery grant) [RGPIN-2018-05089], CRC Tier 2 [NSERC/CRSNG-950-231904], Canada Foundation for Innovation and Québec Ministère de l'Économie, de la Science et de l'Innovation (#37415), and FRQNT Research Support for New Academics to A.H.K., fellowship support from School of Graduate Studies (SGS), Faculty of Arts and Sciences, Concordia University to M.A., B.M.G and SynBioApps program sponsored by NSERC-CREATE to M.V. E.M.M. acknowledges support from the Welch Foundation (F-1515), Army Research Office (W911NF-12-1-0390), and National Institutes of Health (R35 GM122480). The authors declare no competing interests.

Contributions

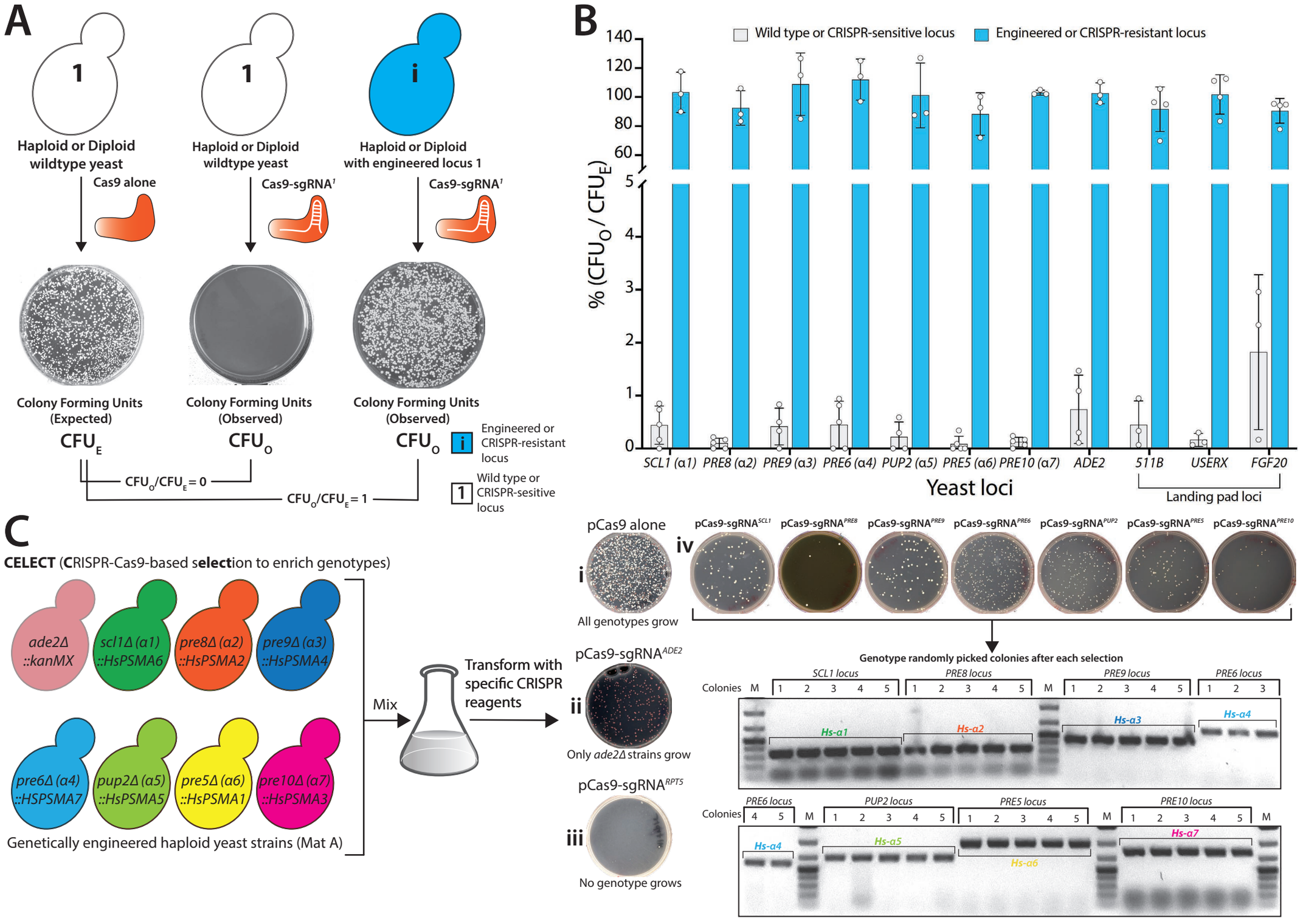
M.A., B.M.G., J.M.L., A.H.K., and E.M.M. designed the research. M.A., B.M.G., J.M.L., and M.V. performed the experiments. A.H.K. and E.M.M. supervised the research. M.A., B.M.G. A.H.K., and E.M.M. wrote the manuscript with editing help from J.M.L., and M.V.

References

1. Botstein, D. & Fink, G. R. Yeast: an experimental organism for 21st Century biology. *Genetics* **189**, 695–704 (2011).
2. Costanzo, M. *et al.* A global genetic interaction network maps a wiring diagram of cellular function. *Science* **353**, aaf1420 (2016).
3. Richardson, S. M. *et al.* Design of a synthetic yeast genome. *Science* **355**, 1040–1044 (2017).
4. Suzuki, Y. *et al.* Bacterial genome reduction using the progressive clustering of deletions via yeast sexual cycling. *Genome Res* **25**, 435–444 (2015).
5. Barbieri, E. M., Muir, P., Akhuetie-Oni, B. O., Yellman, C. M. & Isaacs, F. J. Precise Editing at DNA Replication Forks Enables Multiplex Genome Engineering in Eukaryotes. *Cell* **171**, 1453-1467.e13 (2017).
6. Gibson, D. G. *et al.* Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome. *Science* **319**, 1215–1220 (2008).
7. Fredens, J. *et al.* Total synthesis of Escherichia coli with a recoded genome. *Nature* **569**, 514–518 (2019).
8. Lajoie, M. J. *et al.* Genomically recoded organisms expand biological functions. *Science* **342**, 357–360 (2013).
9. Truong, D. M. & Boeke, J. D. Resetting the Yeast Epigenome with Human Nucleosomes. *Cell* **171**, 1508-1519.e13 (2017).
10. Agmon, N. *et al.* Phylogenetic debugging of a complete human biosynthetic pathway transplanted into yeast. *Nucleic Acids Res* **48**, 486–499 (2020).
11. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **8**, 2281–2308 (2013).
12. DiCarlo, J. E. *et al.* Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. *Nucleic Acids Res* **41**, 4336–4343 (2013).
13. Akhmetov, A. *et al.* Single-step Precision Genome Editing in Yeast Using CRISPR-Cas9. *Bio Protoc* **8**, (2018).
14. Kachroo, A. H. *et al.* Systematic bacterialization of yeast genes identifies a near-universally swappable pathway. *Elife* **6**, (2017).
15. Lee, M. E., DeLoache, W. C., Cervantes, B. & Dueber, J. E. A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synth Biol* **4**, 975–986 (2015).
16. Kuzmin, E., Costanzo, M., Andrews, B. & Boone, C. Synthetic Genetic Array Analysis. *Cold Spring Harb Protoc* **2016**, pdb.prot088807 (2016).
17. Suzuki, Y. *et al.* Knocking out multi-gene redundancies via cycles of sexual assortment and fluorescence selection. *Nat Methods* **8**, 159–164 (2011).
18. Labunskyy, V. M. *et al.* The insertion Green Monster (iGM) method for expression of multiple exogenous genes in yeast. *G3 (Bethesda)* **4**, 1183–1191 (2014).
19. Xu, H. *et al.* Chromosome drives via CRISPR-Cas9 in yeast. *Nat Commun* **11**, 4344 (2020).
20. Yan, Y. & Finnigan, G. C. Development of a multi-locus CRISPR gene drive system in budding yeast. *Sci Rep* **8**, 17277 (2018).
21. DiCarlo, J. E., Chavez, A., Dietz, S. L., Esvelt, K. M. & Church, G. M. Safeguarding CRISPR-Cas9 gene drives in yeast. *Nat Biotechnol* **33**, 1250–1255 (2015).

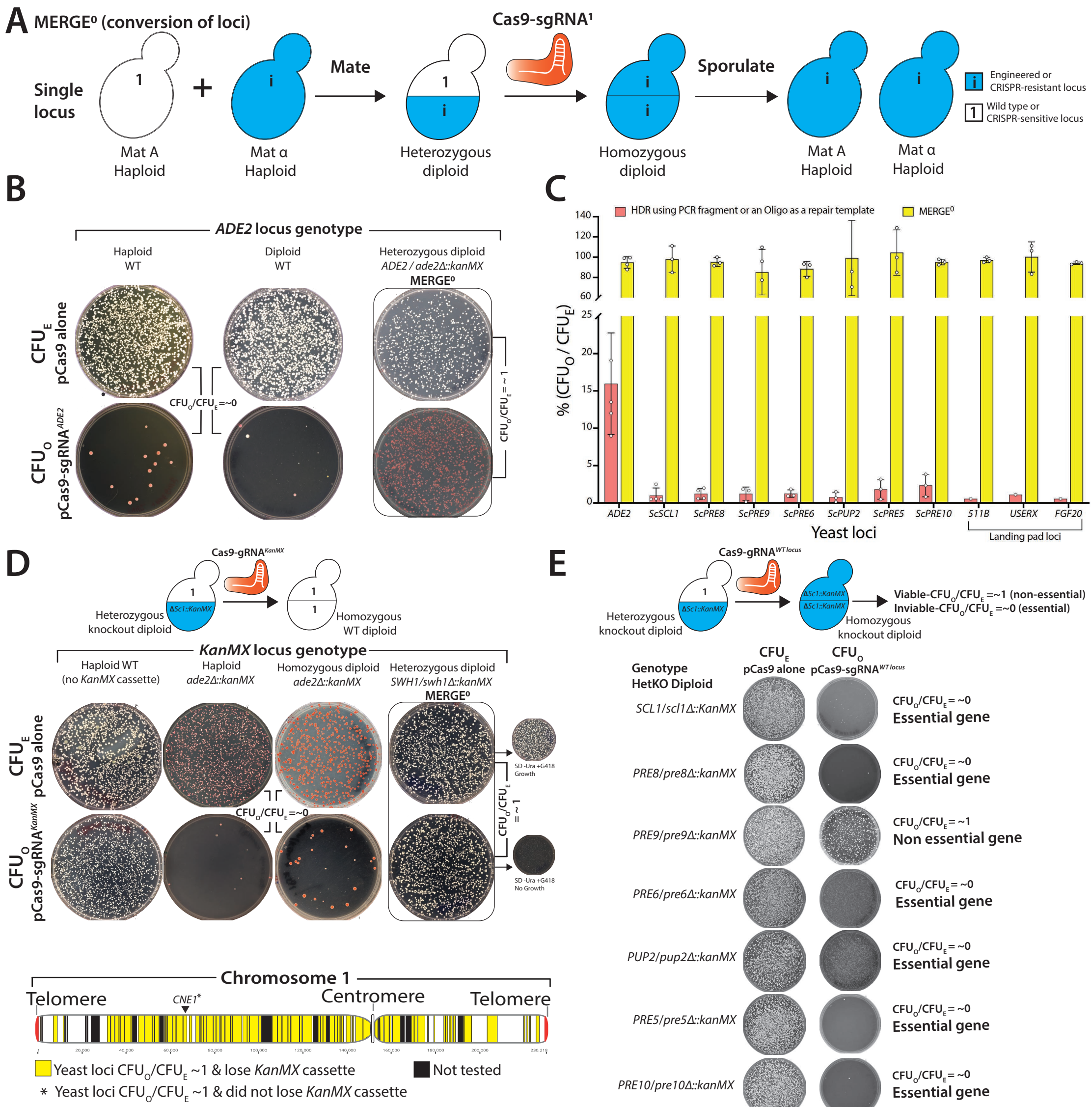
362 22. Roggenkamp, E. *et al.* CRISPR-UnLOCK: Multipurpose Cas9-Based Strategies for
363 Conversion of Yeast Libraries and Strains. *Front Microbiol* **8**, 1773 (2017).
364 23. Kramer, K. M., Brock, J. A., Bloom, K., Moore, J. K. & Haber, J. E. Two different types of
365 double-strand breaks in Saccharomyces cerevisiae are repaired by similar RAD52-independent,
366 nonhomologous recombination events. *Mol Cell Biol* **14**, 1293–1301 (1994).
367 24. Kachroo, A. H. *et al.* Evolution. Systematic humanization of yeast genes reveals
368 conserved functions and genetic modularity. *Science* **348**, 921–925 (2015).
369 25. Winzeler, E. A. *et al.* Functional characterization of the S. cerevisiae genome by gene
370 deletion and parallel analysis. *Science* **285**, 901–906 (1999).
371 26. Pan, X. *et al.* A robust toolkit for functional profiling of the yeast genome. *Mol Cell* **16**,
372 487–496 (2004).
373 27. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**,
374 819–823 (2013).
375 28. Budenholzer, L., Cheng, C. L., Li, Y. & Hochstrasser, M. Proteasome Structure and
376 Assembly. *J Mol Biol* **429**, 3500–3524 (2017).
377 29. Hirano, Y. *et al.* Cooperation of multiple chaperones required for the assembly of
378 mammalian 20S proteasomes. *Mol Cell* **24**, 977–984 (2006).
379 30. Le Tallec, B. *et al.* 20S proteasome assembly is orchestrated by two distinct pairs of
380 chaperones in yeast and in mammals. *Mol Cell* **27**, 660–674 (2007).
381 31. Schnell, H. M. *et al.* Structures of chaperone-associated assembly intermediates reveal
382 coordinated mechanisms of proteasome biogenesis. *Nat Struct Mol Biol* **28**, 418–425 (2021).
383 32. Morris, E. P. & da Fonseca, P. C. A. How to build a proteasome. *Nat Struct Mol Biol* **28**,
384 409–410 (2021).
385 33. Stadtmueller, B. M. *et al.* Structure of a proteasome Pba1-Pba2 complex: implications
386 for proteasome assembly, activation, and biological function. *J Biol Chem* **287**, 37371–37382
387 (2012).
388 34. Murata, S., Yashiroda, H. & Tanaka, K. Molecular mechanisms of proteasome assembly.
389 *Nat Rev Mol Cell Biol* **10**, 104–115 (2009).
390 35. Tomko, R. J. & Hochstrasser, M. Molecular architecture and assembly of the eukaryotic
391 proteasome. *Annu Rev Biochem* **82**, 415–445 (2013).
392
393

Figure 1


A
1
1
i
Haploid or Diploid wildtype yeast
Haploid or Diploid wildtype yeast
Haploid or Diploid with engineered locus 1
Cas9 alone
Cas9-sgRNA1
Cas9-sgRNA1
Colony Forming Units (Expected)
CFUE
Colony Forming Units (Observed)
CFUO
Colony Forming Units (Observed)
CFUO
i Engineered or CRISPR-resistant locus
1 Wild type or CRISPR-sesitive locus
CFUO/CFUE = 0
CFUO/CFUE = 1
B
Wild type or CRISPR-sensitive locus
Engineered or CRISPR-resistant locus
% (CFUO / CFUE)
140
120
100
80
60
5
4
3
2
1
0
SCL1 (α1)
PRE8 (α2)
PRE9 (α3)
PRE6 (α4)
PUP2 (α5)
PRE5 (α6)
PRE10 (α7)
ADE2
511B
USERX
FGF20
Yeast loci
Landing pad loci
C
CELECT (CRISPR-Cas9-based selection to enrich genotypes)
ade2Δ ::kanMX
scl1Δ (α1) ::HsPSMA6
pre8Δ (α2) ::HsPSMA2
pre9Δ (α3) ::HsPSMA4
pre6Δ (α4) ::HSPSMA7
pup2Δ (α5) ::HsPSMA5
pre5Δ (α6) ::HsPSMA1
pre10Δ (α7) ::HsPSMA3
Genetically engineered haploid yeast strains (Mat A)
Mix
Transform with specific CRISPR reagents
pCas9 alone
i
All genotypes grow
pCas9-sgRNAADE2
ii
Only ade2Δ strains grow
pCas9-sgRNARPT5
iii
No genotype grows
iv
pCas9-sgRNASCL1
pCas9-sgRNAPRE8
pCas9-sgRNAPRE9
pCas9-sgRNAPRE6
pCas9-sgRNAPUP2
pCas9-sgRNAPRE5
pCas9-sgRNAPRE10
Genotype randomly picked colonies after each selection
SCL1 locus
PRE8 locus
PRE9 locus
PRE6 locus
Colonies
M
Hs-α1
Hs-α2
Hs-α3
Hs-α4
PRE6 locus
PUP2 locus
PRE5 locus
PRE10 locus
Hs-α4
Hs-α5
Hs-α6
Hs-α7

394 **Figure 1. CRISPR-mediated selection (CELECT) enables the enrichment of unique**
395 **genotypes.**
396 **(A) pCas9-sgRNA^{1}** targeted to any yeast locus (locus **1**) leads to lethality (**CFU_O**) compared to
397 the vector without sgRNA (pCas9 alone; **CFU_E**). However, the modification of locus **1** to **i** prevents
398 further targeting by the corresponding CRISPR reagent (**CFU_O**). A simple readout of **CFU_O/CFU_E**
399 enables the identification of modified genotypes. **(B)** Each CRISPR reagent targeted to several
400 wild-type yeast loci shows near 100% lethality (Gray bars or CRISPR-sensitive loci). However,
401 after editing the loci (by generating single-humanized yeast α proteasome genes, *ade2Δ::kanMX*
402 and inserting carotenoid genes at landing pad loci), the strains show resistance to the
403 corresponding CRISPR reagent (Blue bars or CRISPR-resistant loci, **CFU_O/CFU_E = ~1**). **(C)** Each
404 unique genetically modified strain can be enriched from the mixture using the corresponding
405 CRISPR reagent (CELECT). The transformation of pCas9 alone serves as a control (**i**). Haploid
406 *ade2Δ::kanMX* strain quantifies the efficiency of the CRISPR selection (**pCas9-sgRNA^{ADE2}**),
407 enriching only the *ade2Δ* homozygous genotypes (red-colony phenotype) (**ii**). The transformation
408 of **pCas9-sgRNA^{RPT5}** causes lethality as all the strains in the mix harbor a wild-type copy of the
409 *RPT5* gene (**iii**). Similarly, each CRISPR reagent specific to individual yeast α proteasome genes
410 also selected the corresponding humanized strains as demonstrated by PCR-based genotyping
411 of randomly picked colonies (**iv**).

Figure 2


A
MERGE⁰ (conversion of loci)
Cas9-sgRNA¹
Single locus
Mat A Haploid
Mat α Haploid
Mate
Heterozygous diploid
Homozygous diploid
Sporulate
Mat A Haploid
Mat α Haploid
Engineered or CRISPR-resistant locus
Wild type or CRISPR-sensitive locus
B
ADE2 locus genotype
Haploid WT
Diploid WT
Heterozygous diploid ADE2 / ade2Δ::kanMX
MERGE⁰
CFU_E pCas9 alone
CFU_O pCas9-sgRNA^ADE2
CFU_O/CFU_E =~0
CFU_O/CFU_E = ~ 1
C
HDR using PCR fragment or an Oligo as a repair template
MERGE⁰
% (CFU_O / CFU_E)
ADE2
ScSCL1
ScPRE8
ScPRE9
ScPRE6
ScPUP2
ScPRE5
ScPRE10
511B
USERX
FGF20
Yeast loci
Landing pad loci
D
Cas9-gRNA^KanMX
Heterozygous knockout diploid
Homozygous WT diploid
KanMX locus genotype
Haploid WT (no KanMX cassette)
Haploid ade2Δ::kanMX
Homozygous diploid ade2Δ::kanMX
Heterozygous diploid SWH1/swh1Δ::kanMX
MERGE⁰
CFU_E pCas9 alone
CFU_O pCas9-sgRNA^KanMX
CFU_O/CFU_E =~0
CFU_O/CFU_E = ~ 1
SD -Ura +G418 Growth
SD -Ura +G418 No Growth
Chromosome 1
Telomere
CNE1*
Centromere
Telomere
Yeast loci CFU_O/CFU_E ~1 & lose KanMX cassette
Not tested
* Yeast loci CFU_O/CFU_E ~1 & did not lose KanMX cassette
E
Cas9-gRNA^WT locus
Heterozygous knockout diploid
Homozygous knockout diploid
Viable-CFU_O/CFU_E =~1 (non-essential)
Inviable-CFU_O/CFU_E =~0 (essential)
Genotype HetKO Diploid
CFU_E pCas9 alone
CFU_O pCas9-sgRNA^WT locus
SCL1/scl1Δ::KanMX
CFU_O/CFU_E = ~0
Essential gene
PRE8/pre8Δ::kanMX
CFU_O/CFU_E = ~0
Essential gene
PRE9/pre9Δ::kanMX
CFU_O/CFU_E = ~1
Non essential gene
PRE6/pre6Δ::kanMX
CFU_O/CFU_E = ~0
Essential gene
PUP2/pup2Δ::kanMX
CFU_O/CFU_E = ~0
Essential gene
PRE5/pre5Δ::kanMX
CFU_O/CFU_E = ~0
Essential gene
PRE10/pre10Δ::kanMX
CFU_O/CFU_E = ~0
Essential gene

412 **Figure 2. MERGE level 0 (MERGE^0) efficiently converts a single heterozygous to a**
413 **homozygous locus.**
414 **(A)** Mating haploid yeast strains, each harboring a different allele at a single locus, enables the
415 combination as a heterozygote. CRISPR reagent targeted to one of the alleles (CRISPR-sensitive
416 locus) initiates recombination using the homologous chromosome with a CRISPR-resistant locus
417 as a repair template. Sporulation of the resulting homozygous diploid yields the desired genotype
418 in both mating types. **(B)** CRISPR reagent (**pCas9-sgRNA^{ADE2}**) targeted to the wild type *ADE2*
419 locus quantifies the efficiency of **MERGE^0**. The transformation of **pCas9-sgRNA^{ADE2}** in the wild
420 type haploid or diploid strains is lethal, with few surviving red colonies suggesting efficient ON-
421 target activity (**CFU_O/CFU_E = ~0**). However, the transformation of **pCas9-sgRNA^{ADE2}** in
422 heterozygous diploid *ADE2/ade2Δ::kanMX* strain shows no lethality (**CFU_O/CFU_E = ~1**). Instead,
423 all surviving red colonies suggest efficient conversion to the knockout locus. **(C)** Similarly,
424 CRISPR reagents targeted to several wild-type yeast loci show near 100% lethality with few
425 surviving colonies while simultaneously providing exogenous repair templates (oligo for *ADE2*
426 locus and PCR fragments for the remaining loci) (Red bars). However, each heterozygous diploid
427 displays resistance to CRISPR reagents respectively (**CFU_O/CFU_E = ~1,** Yellow bars). PCR
428 confirmation verified the conversion to the CRISPR-resistant locus. **(D)** CRISPR reagent targeting
429 *KanMX* cassette (**pCas9-sgRNA^{KanMX}**) shows no OFF-target activity in a strain lacking the
430 cassette (**CFU_O/CFU_E = ~1**). However, haploid or diploid strains harboring *KanMX* cassettes as
431 the only allele show lethality (**CFU_O/CFU_E = ~0**), suggesting ON-target activity. However, diploid
432 strain heterozygous for *KanMX* allele (*SWH1*/*swh1Δ::kanMX*) is viable (**CFU_O/CFU_E = ~1**),
433 suggesting conversion to the wild type allele as demonstrated by the loss of G418 resistance after
434 performing **MERGE^0**. Using heterozygous diploid knockout strains (hetKO) arrayed across the
435 entire yeast **chromosome I**, **MERGE^0** similarly converted every knockout allele to the wild type
436 locus (Yellow = **CFU_O/CFU_E = ~1**), except in the case of *CNE1* that showed resistance to the
437 CRISPR reagent but did not lose the *KanMX* cassette. **(E)** Alternatively, using hetKO strains and
438 targeting CRISPR reagent to the wild-type yeast locus enables a single-step gene essentiality
439 assay in yeast. The remarkably high efficiency of **MERGE^0** converts every heterozygous yeast
440 locus to homozygous null, therefore, allowing viability only if the gene is non-essential (as in the
441 case of yeast proteasome α3 gene). All other essential loci show lethality after **MERGE^0**.

Figure 3

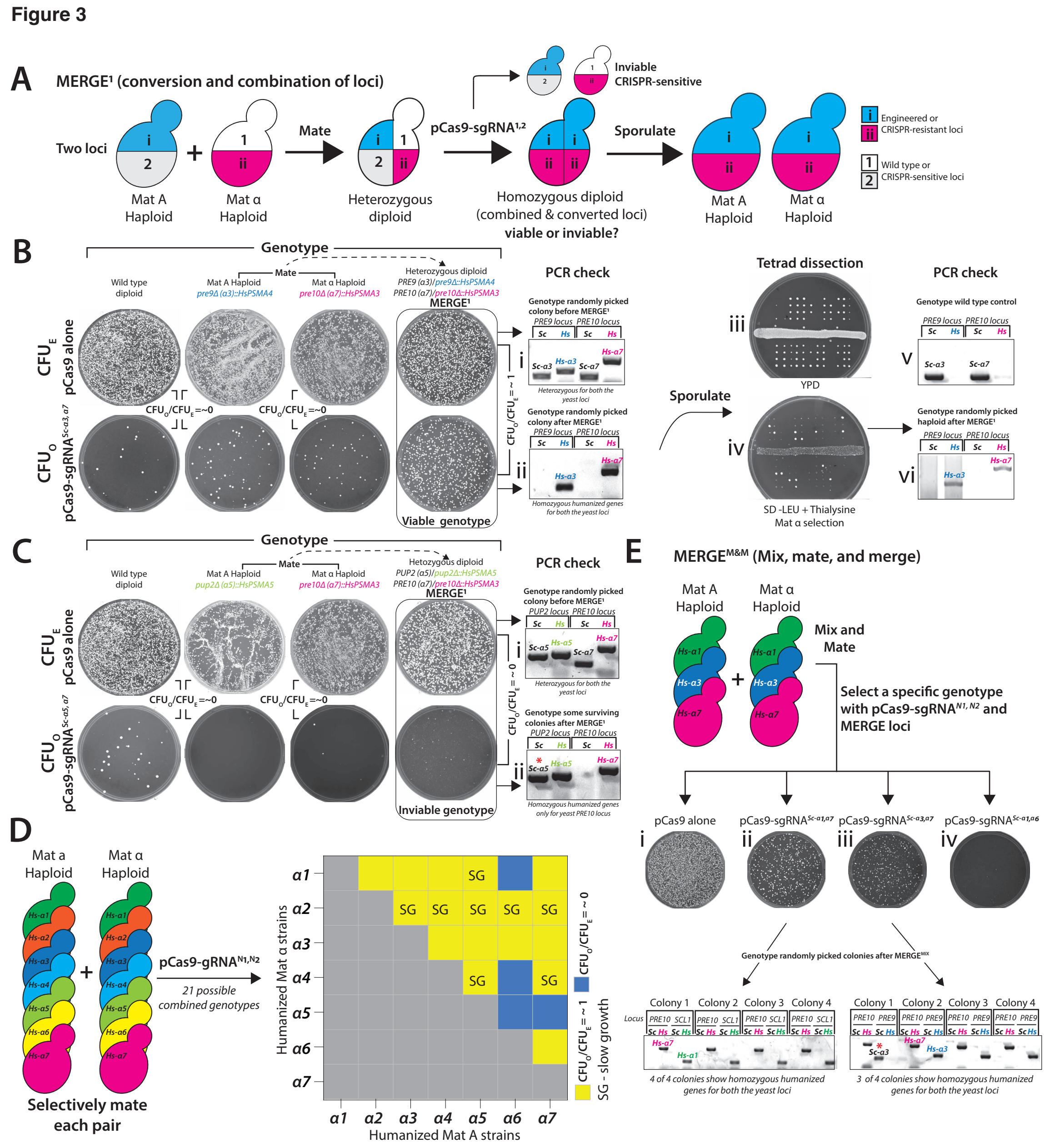
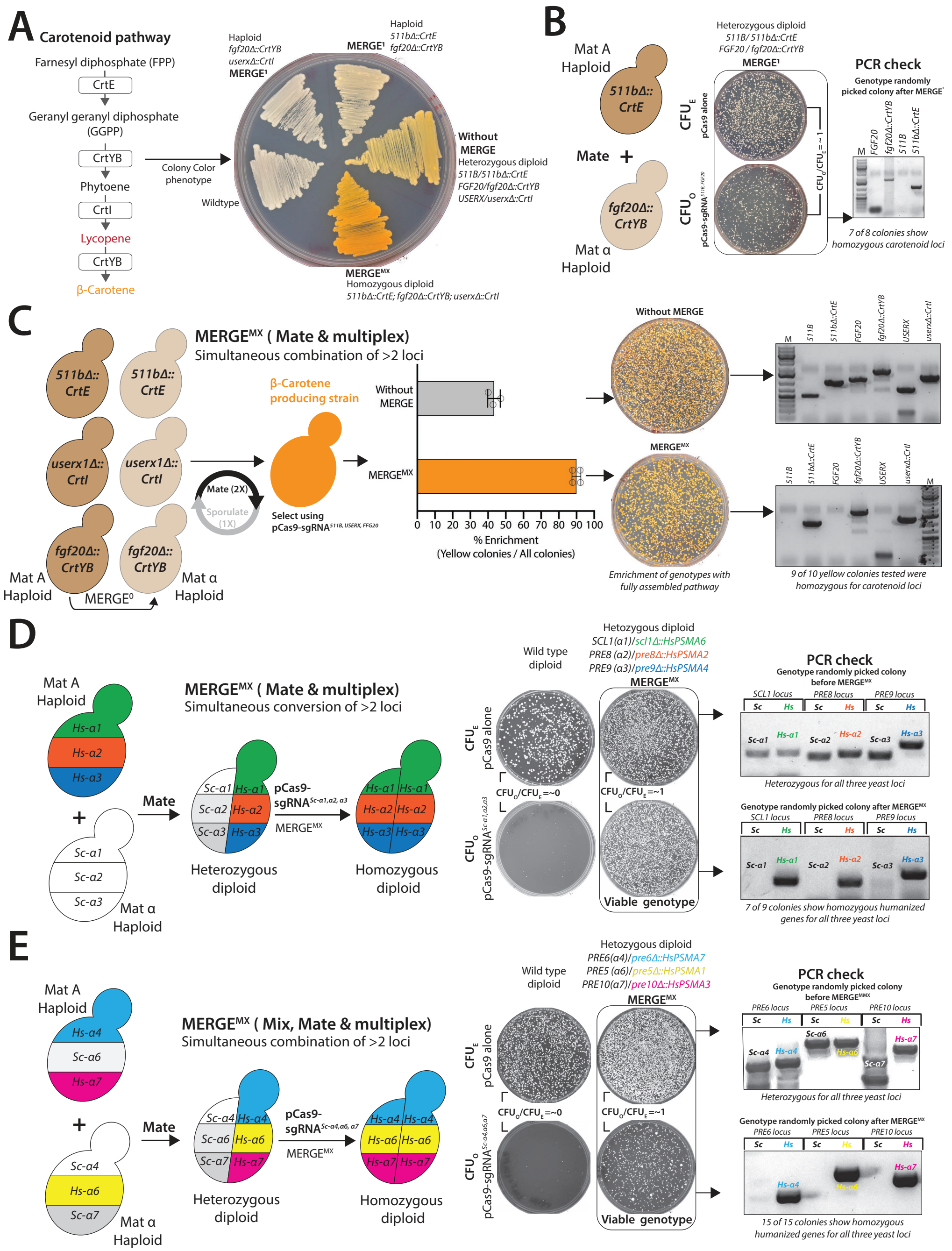

A
MERGE¹ (conversion and combination of loci)
Two loci
Mat A Haploid
Mat α Haploid
Mate
Heterozygous diploid
pCas9-sgRNA¹,²
Inviable CRISPR-sensitive
Homozygous diploid (combined & converted loci)
viable or inviable?
Sporulate
Mat A Haploid
Mat α Haploid
Engineered or CRISPR-resistant loci
Wild type or CRISPR-sensitive loci
B
Genotype
Wild type diploid
Mat A Haploid pre9Δ (a3)::HsPSMA4
Mat α Haploid pre10Δ (a7)::HsPSMA3
Heterozygous diploid PRE9 (a3)/pre9Δ::HsPSMA4 PRE10 (a7)/pre10Δ::HsPSMA3
MERGE¹
CFU_E pCas9 alone
CFU_O pCas9-sgRNA^Sc-a3, a7
CFU_O/CFU_E=~0
CFU_O/CFU_E=~1
Viable genotype
PCR check
Genotype randomly picked colony before MERGE¹
Heterozygous for both the yeast loci
Genotype randomly picked colony after MERGE¹
Homozygous humanized genes for both the yeast loci
Tetrad dissection
YPD
SD -LEU + Thialysine Mat α selection
Genotype wild type control
Genotype randomly picked haploid after MERGE¹
C
Wild type diploid
Mat A Haploid pup2Δ (a5)::HsPSMA5
Mat α Haploid pre10Δ (a7)::HsPSMA3
Hetozygous diploid PUP2 (a5)/pup2Δ::HsPSMA5 PRE10 (a7)/pre10Δ::HsPSMA3
CFU_E pCas9 alone
CFU_O pCas9-sgRNA^Sc-a5, a7
CFU_O/CFU_E=~0
CFU_O/CFU_E=~1
Inviable genotype
Genotype randomly picked colony before MERGE¹
Heterozygous for both the yeast loci
Genotype some surviving colonies after MERGE¹
Homozygous humanized genes only for yeast PRE10 locus
D
Mat a Haploid
Mat α Haploid
pCas9-sgRNA^N1,N2
21 possible combined genotypes
Selectively mate each pair
Humanized Mat α strains
Humanized Mat A strains
SG
CFU_O/CFU_E=~0
CFU_O/CFU_E=~1
SG - slow growth
E
MERGE^M&M (Mix, mate, and merge)
Mat A Haploid
Mat α Haploid
Mix and Mate
Select a specific genotype with pCas9-sgRNA^N1,N2 and MERGE loci
pCas9 alone
pCas9-sgRNA^Sc-a1,a7
pCas9-sgRNA^Sc-a3,a7
pCas9-sgRNA^Sc-a1,a6
Genotype randomly picked colonies after MERGE^MIX
Colony 1
Colony 2
Colony 3
Colony 4
4 of 4 colonies show homozygous humanized genes for both the yeast loci
3 of 4 colonies show homozygous humanized genes for both the yeast loci

Figure 3. MERGE level 1 (MERGE^1) efficiently converts and combines two heterozygous to homozygous loci.

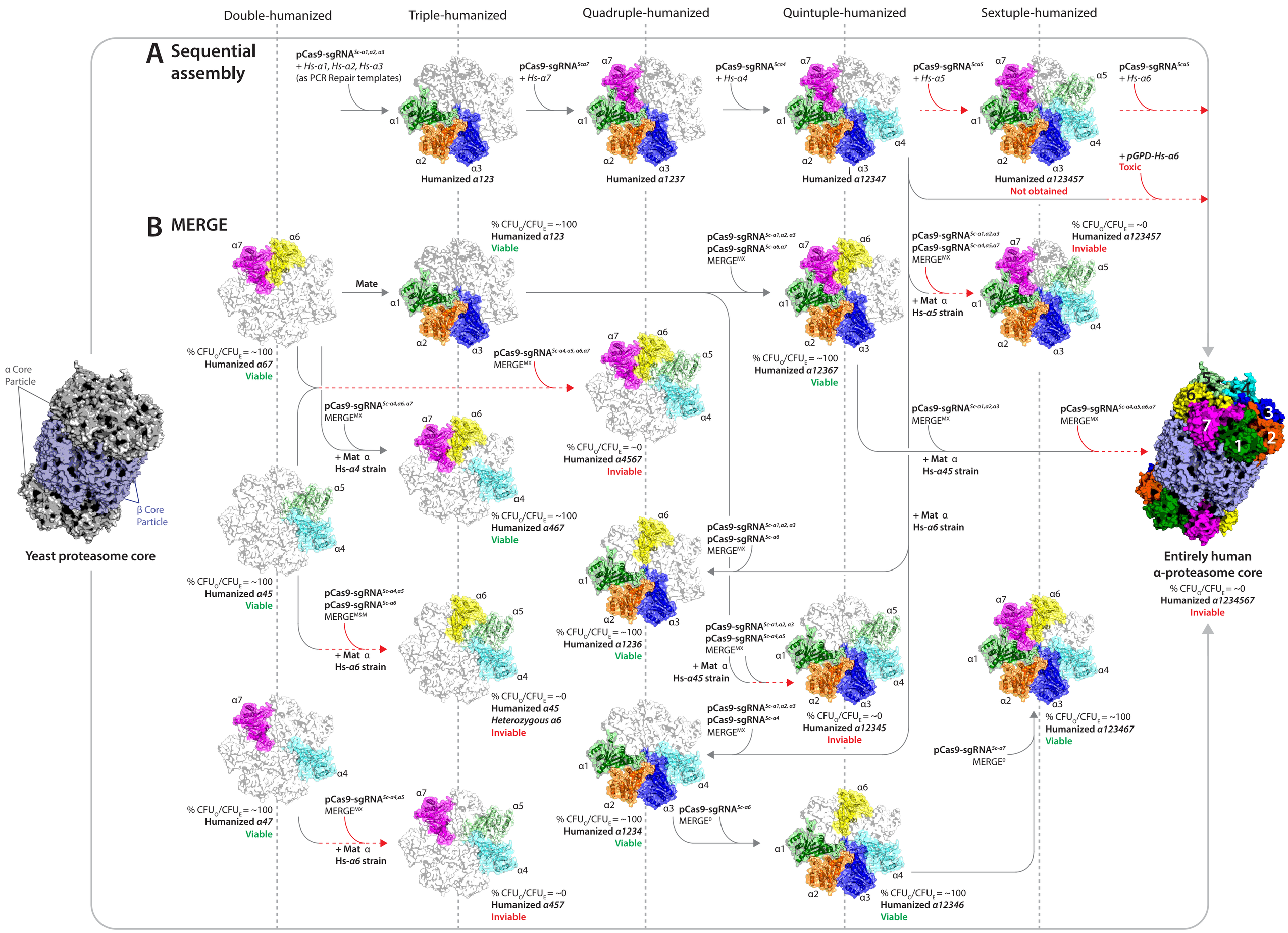
(A) The schematic shows the mating of haploid yeast strains, each harboring two different alleles, enabling the combination as a heterozygote for each locus. Double-sgRNA CRISPR reagent targeted to both the alleles (CRISPR-sensitive loci, **1 & 2**) initiates recombination at both loci using the homologous chromosome with CRISPR-resistant loci (**i & ii**) as a repair template, thus, enabling the simultaneous combination of 2 loci. **CFU_O/CFU_E = ~1** suggests viable combined genotypes. Sporulation of the resulting homozygous diploid yields the desired combined genotype in both mating types. **(B)** Double-sgRNA CRISPR reagent (**pCas9-sgRNA$^{Sc\alpha3,\alpha7}$**) targeted to two wild type yeast loci (proteasome α3 and α7 genes) quantifies the efficiency of **MERGE^1**. The transformation of **pCas9-sgRNA$^{Sc\alpha3,\alpha7}$** in the wild type, and single-humanized haploid Hsα3 or Hsα7 strains is lethal, with few surviving colonies (**CFU_O/CFU_E = ~0**). However, the transformation of **pCas9-sgRNA$^{Sc\alpha3,\alpha7}$** in a heterozygous diploid humanized strain shows no lethality (**CFU_O/CFU_E = ~1**). PCR-based genotyping of surviving diploid strains after **MERGE^1** shows conversion of both yeast to the humanized loci (**ii**) compared to before **MERGE^1** (**i**). Sporulation generated the paired-humanized haploids in both mating-types (shown here as Tetrads spotted on YPD (**iii**) and Mat α selection-**iv**). PCR check of haploid strains confirms the combination of humanized loci (**vi**). PCR of wild-type yeast loci is shown as a reference (**v**). **(C)** Alternatively, double-sgRNA CRISPR reagent (**pCas9-sgRNA$^{Sc\alpha5,\alpha7}$**) targeting two wild type yeast loci (proteasome α5 and α7 genes) in a corresponding heterozygous diploid humanized strain showed lethality, suggesting incompatible combination (**CFU_O/CFU_E = ~0**). The transformation of **pCas9-sgRNA$^{Sc\alpha5,\alpha7}$** in the wild type, and single-humanized Hsα*5* or Hsα*7* strains is lethal, with few surviving colonies, suggesting ON-target activity (**CFU_O/CFU_E = ~0**). PCR-based genotyping of surviving colonies after **MERGE^1** shows the conversion of only one (α*5*) yeast to the humanized locus (**ii**) compared to before **MERGE^1** showing heterozygous genotype at both loci (**i**). **(D) MERGE^0** generated 7 humanized α proteasome strains in each mating-type. Mating each single-humanized strain allows a systematic test for every viable double-humanized genotype (21 combined genotypes). Transformation of double-sgRNA CRISPR reagent uniquely targeting yeast genes in each paired heterozygote facilitates the combination of two humanized loci as homozygotes (**CFU_O/CFU_E = ~1, yellow**) while also revealing genotypes that are not permitted (**CFU_O/CFU_E = ~0, blue**). **(E)** Mixing and mating singly humanized genotypes (Hsα1, Hsα3 and Hsα7) allows a random combination of two humanized alleles after double-sgRNA CRISPR selection referred to as mix, mate, and MERGE (**MERGE$^{M\&M}$**). Transformation of **pCas9 alone** serves as a control (**i**). Double-sgRNA CRISPR reagents, **pCas9-sgRNA$^{Sc\alpha1,\alpha7}$** (**ii**) and **pCas9-sgRNA$^{Sc\alpha3,\alpha7}$** (**iii**), specifically enriched the corresponding double-humanized genotype while also converting yeast to humanized loci. PCR-based genotyping of randomly picked colonies confirms *Hsα1α7* (4 of 4 colonies tested) and *Hsα3α7* (3 of 4 colonies tested). Comparatively, the transformation of double-sgRNA CRISPR reagent, **pCas9-sgRNA$^{Sc\alpha1,\alpha6}$** (**iv**), shows no surviving colonies as the genotype does not exist in the mixture.

Figure 4


A
Carotenoid pathway
Farnesyl diphosphate (FPP)
CrtE
Geranyl geranyl diphosphate (GGPP)
CrtYB
Phytoene
CrtI
Lycopene
CrtYB
β-Carotene
Colony Color phenotype
Haploid fgf20Δ::CrtYB userxΔ::CrtI MERGE¹
Haploid 511bΔ::CrtE fgf20Δ::CrtYB MERGE¹
Wildtype
Without MERGE
Heterozygous diploid 511B/511bΔ::CrtE FGF20/fgf20Δ::CrtYB USERX/userxΔ::CrtI
MERGEMX
Homozygous diploid 511bΔ::CrtE; fgf20Δ::CrtYB; userxΔ::CrtI
B
Mat A Haploid
511bΔ:: CrtE
Mate +
fgf20Δ:: CrtYB
Mat α Haploid
Heterozygous diploid 511B/ 511bΔ::CrtE FGF20 / fgf20Δ::CrtYB
MERGE¹
CFUE pCas9 alone
CFUO pCas9-sgRNA511B, FGF20
CFUO/CFUE = ~ 1
PCR check
Genotype randomly picked colony after MERGE¹
M FGF20 fgf20Δ::CrtYB 511B 511bΔ::CrtE
7 of 8 colonies show homozygous carotenoid loci
C
MERGEMX (Mate & multiplex)
Simultaneous combination of >2 loci
511bΔ:: CrtE
511bΔ:: CrtE
userx1Δ:: CrtI
userx1Δ:: CrtI
fgf20Δ:: CrtYB
fgf20Δ:: CrtYB
Mat A Haploid
Mat α Haploid
MERGE⁰
Mate (2X)
Sporulate (1X)
β-Carotene producing strain
Select using pCas9-sgRNA511B, USERX, FFG20
Without MERGE
MERGEMX
0 10 20 30 40 50 60 70 80 90 100
% Enrichment (Yellow colonies / All colonies)
Without MERGE
MERGEMX
Emrichment of genotypes with fully assembled pathway
M 511B 511bΔ::CrtE FGF20 fgf20Δ::CrtYB USERX userxΔ::CrtI
9 of 10 yellow colonies tested were homozygous for carotenoid loci
D
Mat A Haploid
Hs-α1
Hs-α2
Hs-α3
+
Sc-α1
Sc-α2
Sc-α3
Mat α Haploid
Mate
MERGEMX (Mate & multiplex)
Simultaneous conversion of >2 loci
Sc-α1 Hs-α1
Sc-α2 Hs-α2
Sc-α3 Hs-α3
pCas9-sgRNASc-α1,α2, α3
MERGEMX
Heterozygous diploid
Homozygous diploid
Wild type diploid
Hetozygous diploid SCL1(α1)/scl1Δ::HsPSMA6 PRE8 (α2)/pre8Δ::HsPSMA2 PRE9 (α3)/pre9Δ::HsPSMA4
MERGEMX
CFUE pCas9 alone
CFUO pCas9-sgRNASc-α1,α2,α3
CFUO/CFUE=~0
CFUO/CFUE=~1
Viable genotype
PCR check
Genotype randomly picked colony before MERGEMX
SCL1 locus PRE8 locus PRE9 locus
Sc Hs
Sc-α1 Hs-α1 Sc-α2 Hs-α2 Sc-α3 Hs-α3
Heterozygous for all three yeast loci
Genotype randomly picked colony after MERGEMX
7 of 9 colonies show homozygous humanized genes for all three yeast loci
E
Mat A Haploid
Hs-α4
Sc-α6
Hs-α7
+
Sc-α4
Hs-α6
Sc-α7
Mat α Haploid
Mate
MERGEMX (Mix, Mate & multiplex)
Simultaneous combination of >2 loci
Sc-α4 Hs-α4
Sc-α6 Hs-α6
Sc-α7 Hs-α7
pCas9-sgRNASc-α4,α6, α7
MERGEMX
Heterozygous diploid
Homozygous diploid
Wild type diploid
Hetozygous diploid PRE6(α4)/pre6Δ::HsPSMA7 PRE5 (α6)/pre5Δ::HsPSMA1 PRE10(α7)/pre10Δ::HsPSMA3
MERGEMX
CFUE pCas9 alone
CFUO pCas9-sgRNASc-α4,α6,α7
CFUO/CFUE=~0
CFUO/CFUE=~1
Viable genotype
PCR check
Genotype randomly picked colony before MERGEMMX
PRE6 locus PRE5 locus PRE10 locus
Sc Hs
Sc-α4 Hs-α4 Sc-α6 Hs-α6 Sc-α7 Hs-α7
Heterozygous for all three yeast loci
Genotype randomly picked colony after MERGEMX
15 of 15 colonies show homozygous humanized genes for all three yeast loci

481 **Figure 4. MERGE mate and multiplex (MERGE^{MX}) is scalable to combine multiple**
482 **engineered loci.**
483 **(A)** Schematic shows the carotenoid pathway genes and the metabolic intermediates leading to
484 the color colony phenotype. A complete pathway (*CrtE*, *CrtYB* & *CrtI*) leads to an orange colony
485 appearance, whereas the partial assembly (*CrtYB* & *CrtI*) provides an off-white colony phenotype.
486 **MERGE^{1}** provided single-carotenoid engineered strains in both mating types. The homozygous
487 diploid for a complete carotenoid pathway shows an intense orange colony phenotype than the
488 heterozygous strain. **(B)** Schematic showing **MERGE1** combining two engineered genotypes at
489 landing pad loci. Double-sgRNA CRISPR reagent (**pCas9-sgRNA^{511B,FGF20}**) targeting two landing
490 pad loci (*511B* and *FGF20*) quantifies the efficiency of **MERGE^{1}**. The transformation of **pCas9-**
491 **sgRNA^{511B,FGF20}** in double heterozygous diploid strain shows no lethality (**CFU_O/CFU_E = ~1**). PCR-
492 based genotyping of several colonies after **MERGE^{1}** shows the conversion to the engineered
493 carotenoid loci. **(C)** Single-carotenoid gene strains of both mating-types were mixed, mated (2X)
494 and sporulated (1X). The mixture transformed with **pCas9 alone** shows an equal distribution of
495 white, off-white and orange-colored colonies. However, the triple-sgRNA **pCas9-**
496 **sgRNA^{511B,FGF20,USERX}** transformed mix significantly enriched the yellow colonies. PCR-based
497 genotyping of several colonies showed **MERGE^{MX}** successfully homozygosed all loci, whereas
498 the yellow colonies on the unselected plate showed heterozygous loci. **(D)** Schematic shows
499 mating haploid triple-humanized Hs*a*1,*a*2,*a*3 strain with a wild-type yeast enables the combination
500 of 3 humanized loci as heterozygotes. Triple-sgRNA CRISPR reagent (**pCas9-sgRNA^{Sca1,a2,a3}**)
501 targeted to the corresponding wild-type yeast alleles (CRISPR-sensitive loci) simultaneously
502 initiates recombination at all 3 loci using the homologous chromosome with humanized CRISPR-
503 resistant loci as a repair template. The transformation of **pCas9-sgRNA^{Sca1,a2,a3}** in the wild type is
504 lethal, suggesting ON-target activity (**CFU_O/CFU_E = ~0**). However, the transformation of **pCas9-**
505 **sgRNA^{Sca1,a2,a3}** in a diploid triple heterozygous humanized strain shows no lethality (**CFU_O/CFU_E**
506 **= ~1**). PCR-based genotyping of surviving colonies after **MERGE^{MX}** confirms the conversion of all
507 3 yeast to the humanized loci compared to before **MERGE^{MX}**. The transformation of a plasmid
508 without sgRNA (**pCas9 alone**) serves as a transformation efficiency control (**CFU_E**). **(E)**
509 **MERGE^{MX}** similarly combined >2 loci after mating a strain with 2 humanized loci (Hs*a*4,*a*7, Mat
510 A) with a strain carrying 1 humanized locus (Hs*a*6, Mat α). The transformation of **pCas9-**
511 **sgRNA^{Sca4,a6,a7}** in the wild type is lethal, suggesting ON-target activity (**CFU_O/CFU_E = ~0**).
512 However, the transformation of **pCas9-sgRNA^{Sca4,a6,a7}** in a mated mix selected a diploid triple-
513 humanized strain (**CFU_O/CFU_E = ~1**). PCR-based genotyping of several colonies after **MERGE^{MX}**
514 shows conversion of all yeast to the humanized loci.

Figure 5


Double-humanized
Triple-humanized
Quadruple-humanized
Quintuple-humanized
Sextuple-humanized
A Sequential assembly
pCas9-sgRNA Sc-α1,α2,α3 + Hs-α1, Hs-α2, Hs-α3 (as PCR Repair templates)
Humanized α123
pCas9-sgRNA Sca7 + Hs-α7
Humanized α1237
pCas9-sgRNA Sca4 + Hs-α4
Humanized α12347
pCas9-sgRNA Sca5 + Hs-α5
Humanized α123457
Not obtained
pCas9-sgRNA Sca5 + Hs-α6
+ pGPD-Hs-α6
Toxic
B MERGE
Mate
% CFUO/CFUE = ~100
Humanized α123
Viable
pCas9-sgRNA Sc-α1,α2,α3
pCas9-sgRNA Sc-α6,α7
MERGEMX
% CFUO/CFUE = ~100
Humanized α12367
Viable
pCas9-sgRNA Sc-α1,α2,α3
pCas9-sgRNA Sc-α4,α5,α7
MERGEMX
+ Mat α
Hs-α5 strain
% CFUO/CFUE = ~0
Humanized α123457
Inviable
% CFUO/CFUE = ~100
Humanized α67
Viable
pCas9-sgRNA Sc-α4,α5,α6,α7
MERGEMX
% CFUO/CFUE = ~0
Humanized α4567
Inviable
pCas9-sgRNA Sc-α4,α6,α7
MERGEMX
+ Mat α
Hs-α4 strain
% CFUO/CFUE = ~100
Humanized α467
Viable
pCas9-sgRNA Sc-α1,α2,α3
MERGEMX
+ Mat α
Hs-α45 strain
pCas9-sgRNA Sc-α4,α5,α6,α7
MERGEMX
Entirely human α-proteasome core
% CFUO/CFUE = ~0
Humanized α1234567
Inviable
α Core Particle
β Core Particle
Yeast proteasome core
% CFUO/CFUE = ~100
Humanized α45
Viable
pCas9-sgRNA Sc-α4,α5
pCas9-sgRNA Sc-α6
MERGEM&M
+ Mat α
Hs-α6 strain
% CFUO/CFUE = ~0
Humanized α45
Heterozygous α6
Inviable
pCas9-sgRNA Sc-α1,α2,α3
pCas9-sgRNA Sc-α6
MERGEMX
% CFUO/CFUE = ~100
Humanized α1236
Viable
pCas9-sgRNA Sc-α1,α2,α3
pCas9-sgRNA Sc-α4,α5
MERGEMX
+ Mat α
Hs-α45 strain
% CFUO/CFUE = ~0
Humanized α12345
Inviable
% CFUO/CFUE = ~100
Humanized α123467
Viable
pCas9-sgRNA Sc-α7
MERGE0
% CFUO/CFUE = ~100
Humanized α47
Viable
pCas9-sgRNA Sc-α4,α5
MERGEMX
+ Mat α
Hs-α6 strain
% CFUO/CFUE = ~0
Humanized α457
Inviable
pCas9-sgRNA Sc-α1,α2,α3
pCas9-sgRNA Sc-α4
MERGEMX
% CFUO/CFUE = ~100
Humanized α1234
Viable
pCas9-sgRNA Sc-α6
MERGE0
% CFUO/CFUE = ~100
Humanized α12346
Viable

515 **Figure 5. MERGE reveals a fitness-driven path to the humanization of most of the *α***
516 **proteasome core in yeast.**
517 The schematic shows the transition of heptameric yeast *α* proteasome to humanized *α*
518 proteasome core. **(A)** A sequential strategy used a triple-humanized Hs*α*1,*α*2,*α*3 strain as a
519 background to progressively humanize the rest of yeast *α* core genes by co-transforming a
520 CRISPR reagent and a human gene repair template. Choices to sequentially humanize the *α*
521 proteasome core were made depending on the success of the prior effort. The strategy permitted
522 the humanization of yeast *α*7 followed by α4. Whereas several attempts to humanize yeast *α*5
523 and *α*6 failed. The expression of human Hs*α*6 in a quintuple-humanized Hs*α*1,*α*2*α*3,*α*4,*α*7 strain
524 resulted in a lethal phenotype. **(B) MERGE** provided a clear readout of the fitness of combined
525 genotypes while revealing incompatible combinations of the humanized yeast α proteasome core.
526 Mating distinct humanized yeast combinations (show as connecting lines) followed by **MERGE^{MX}**
527 tested various triple-, quadruple-, quintuple- and sextuple-humanized *α*-proteasome yeast strains.
528 **MERGE** generated viable triple-humanized combinations Hsα4,α5,α6 and Hsα4,α5,α7 (indicated
529 as dashed lines in red), suggesting incompatible combinations. Similarly, while **MERGE^{1}**
530 facilitated the generation of Hs*α*4,*α*5 and Hs*α*6,*α*7 genotypes, the subsequent **MERGE^{MX}** to
531 generate quadrupled Hs*α*4,*α*5,*α*6,*α*7 failed. Using **MERGE^{MX}** followed by **MERGE^{0}** identified
532 several quadruples, quintuples and one sextuple-humanized genotype (Hsα1,α2,α3,α4,α6,α7),
533 revealing a fitness-driven path to the humanization of near entire yeast α-proteasome core.
534 Whereas identifying humanized combinations (>2), comprising the Hsα5 subunit, as incompatible
535 genotypes. Mating of Hsα1,α2,α3,α6,α7 and Hsα4,α5 strains generated a heterozygous human-
536 yeast diploid for all 7 α proteasome core genes. Using multiplex-sgRNA CRISPR reagents
537 permitted verifying if the entire human α-proteasome core is feasible. A triple-sgRNA CRISPR
538 reagent (**pCas9-sgRNA$^{Sc\text{-}\alpha1,\alpha2,\alpha3}$**) homozygosed 3 of 7 humanized loci. However, a subsequent
539 **MERGE^{MX},** using a quadruple-sgRNA CRISPR reagent (**pCas9-sgRNA$^{Sc\text{-}\alpha4,\alpha5,\alpha6,\alpha7}$**), failed to
540 obtain a viable genotype. Proteasome core structures were generated using Pymol and PDB-
541 1RYP. Colored structures show humanized α yeast subunits.

Figure S1

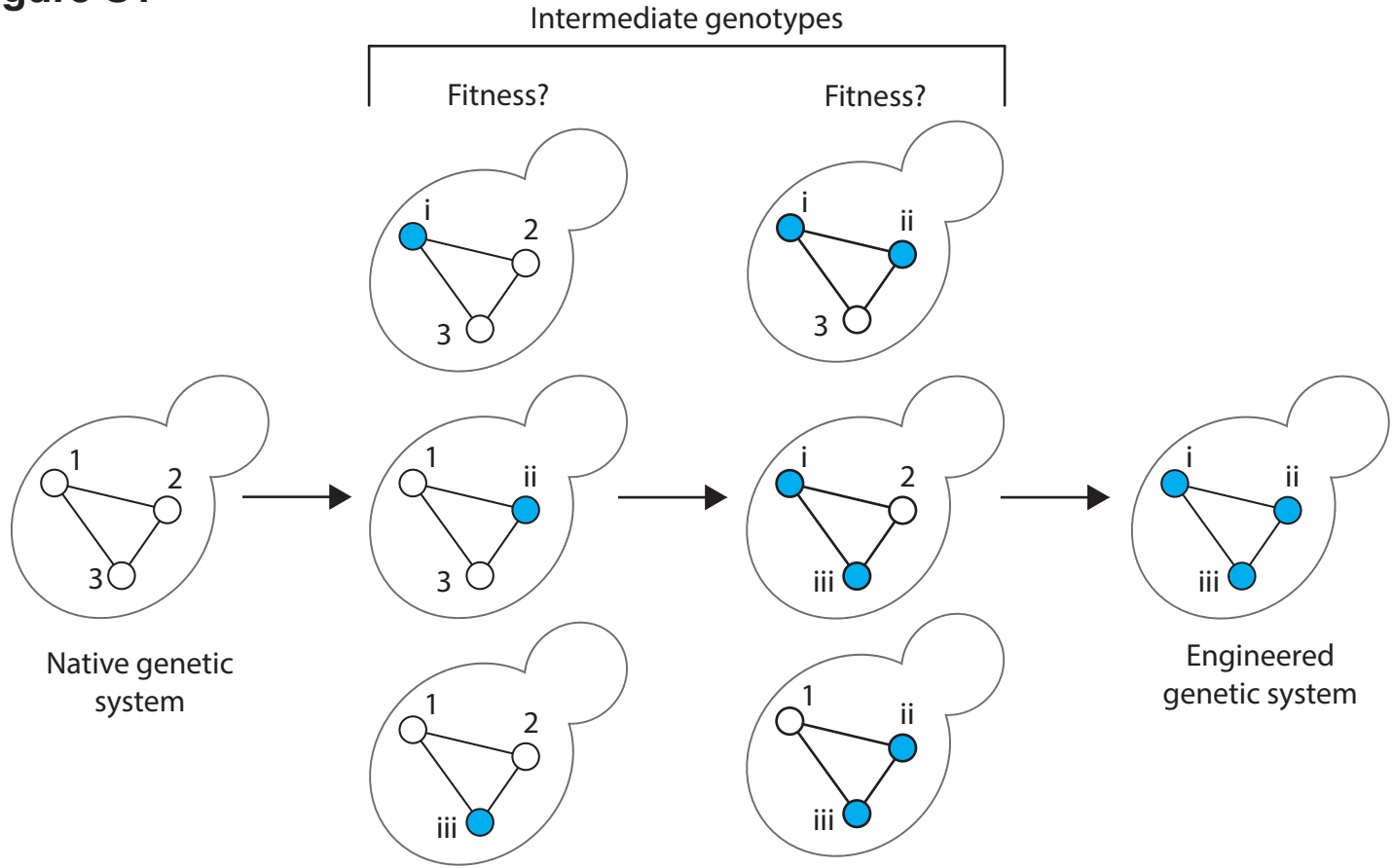
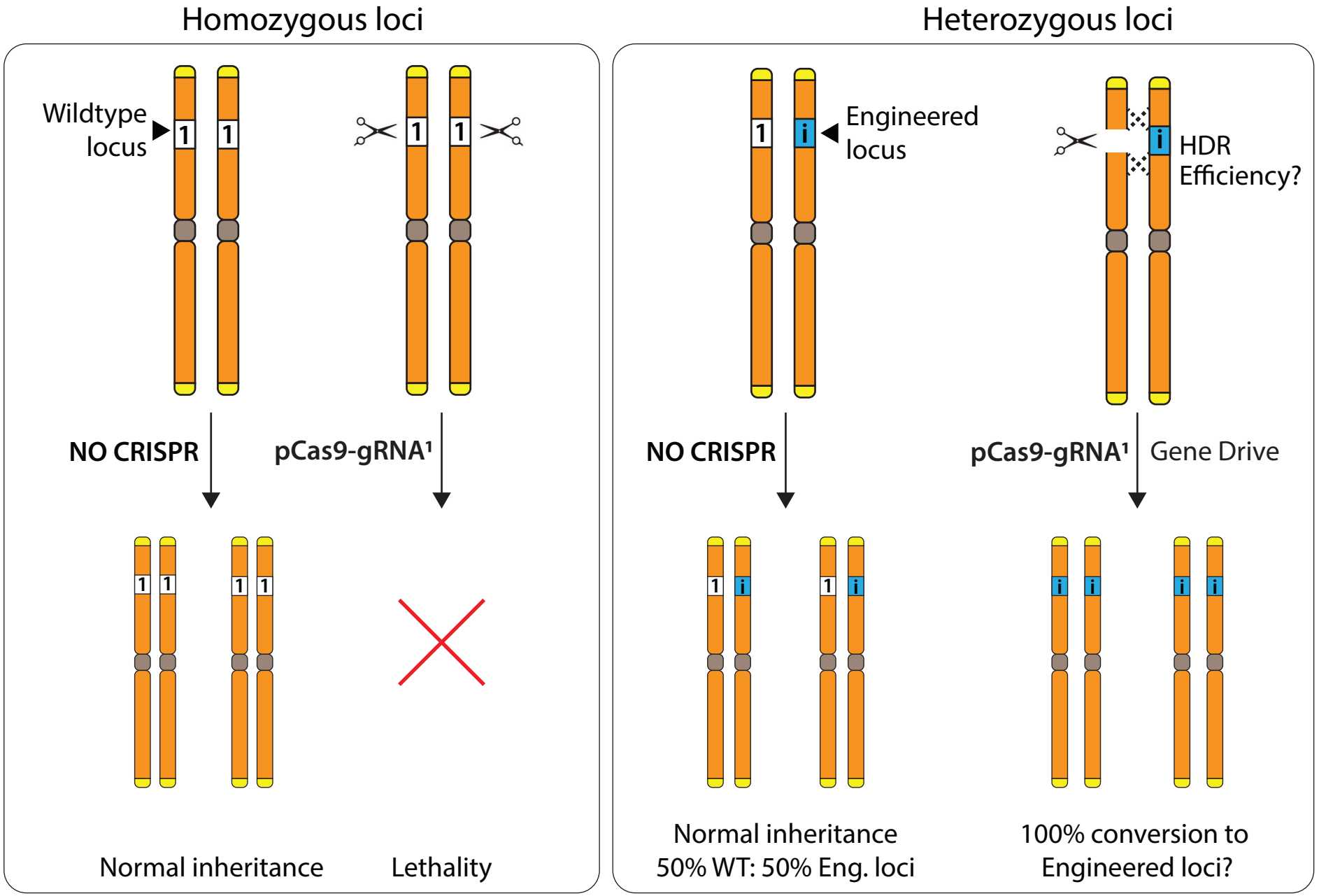

Intermediate genotypes
Fitness?
Fitness?
i
2
3
i
ii
3
1
2
3
1
ii
3
i
2
iii
i
ii
iii
1
2
iii
1
ii
iii
Native genetic system
Engineered genetic system

Figure S1. Multiple routes to the engineering of biological systems using a sequential replacement strategy.

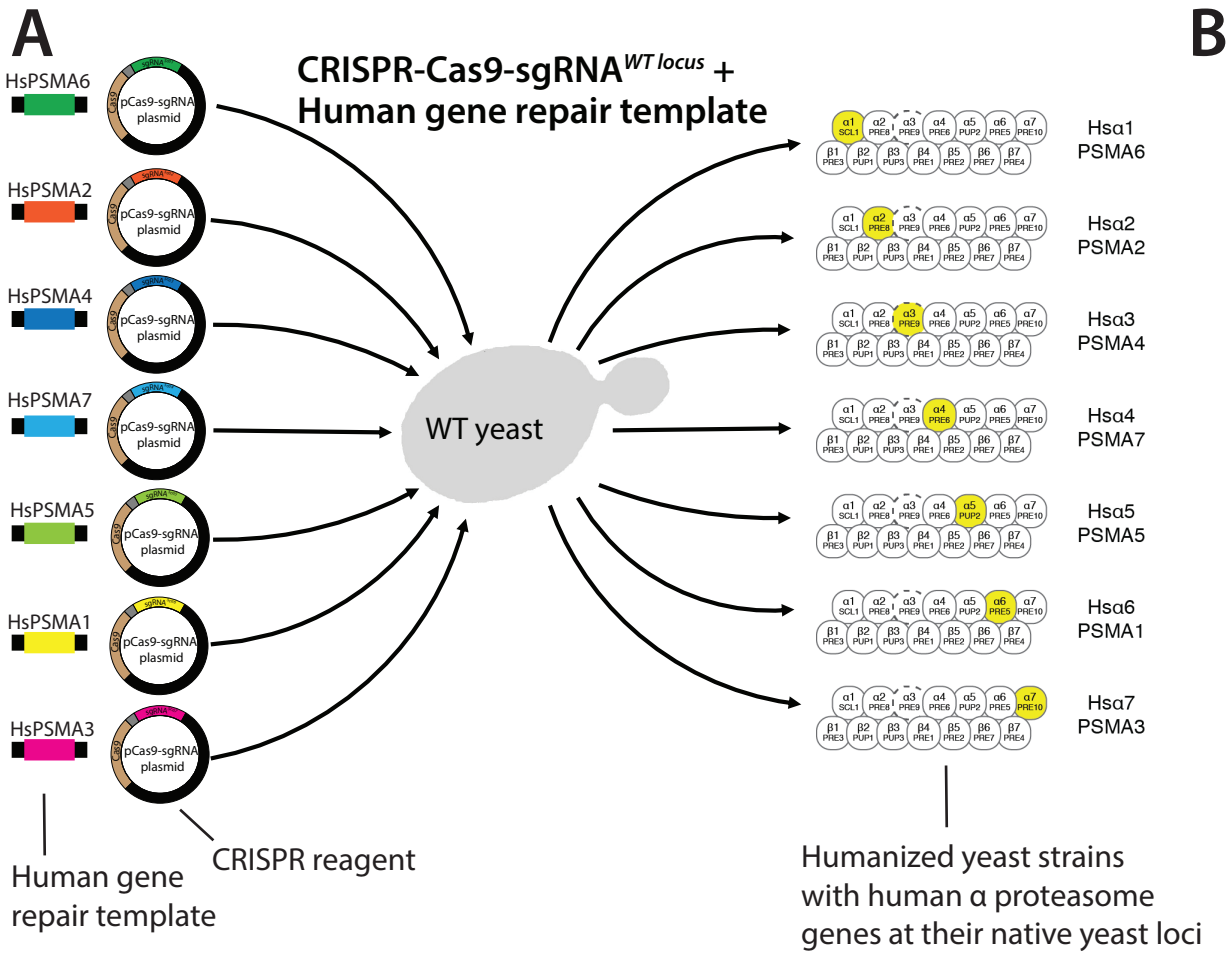
Let's examine a simple 3-gene pathway. A sequential replacement strategy to fully engineer this module will navigate 6 possible intermediate genotypes. The following formula can calculate the number of genotypes: $nCr = (n!/n - r!)\ (r!)$, where n represents a total number of genes to be combined (for example, depicted above, n = 3). And r represents the combined intermediates, like for a single-gene intermediates r = 1 and similarly for a 2-gene intermediates r = 2. Furthermore, while the strategy explores 6 intermediate genotypes, the distinct routes to complete pathway engineering are 9 as some genotypes may require gene 1 and then gene 2, or gene 2 and then gene 1. The higher the number of genes to be engineered, the more the number of intermediate genotypes to be explored. As a reference, for the humanization of the heptameric α proteasome core, a total of 126 intermediate hybrid yeast-human genotypes exist towards humanizing the entire yeast α core.

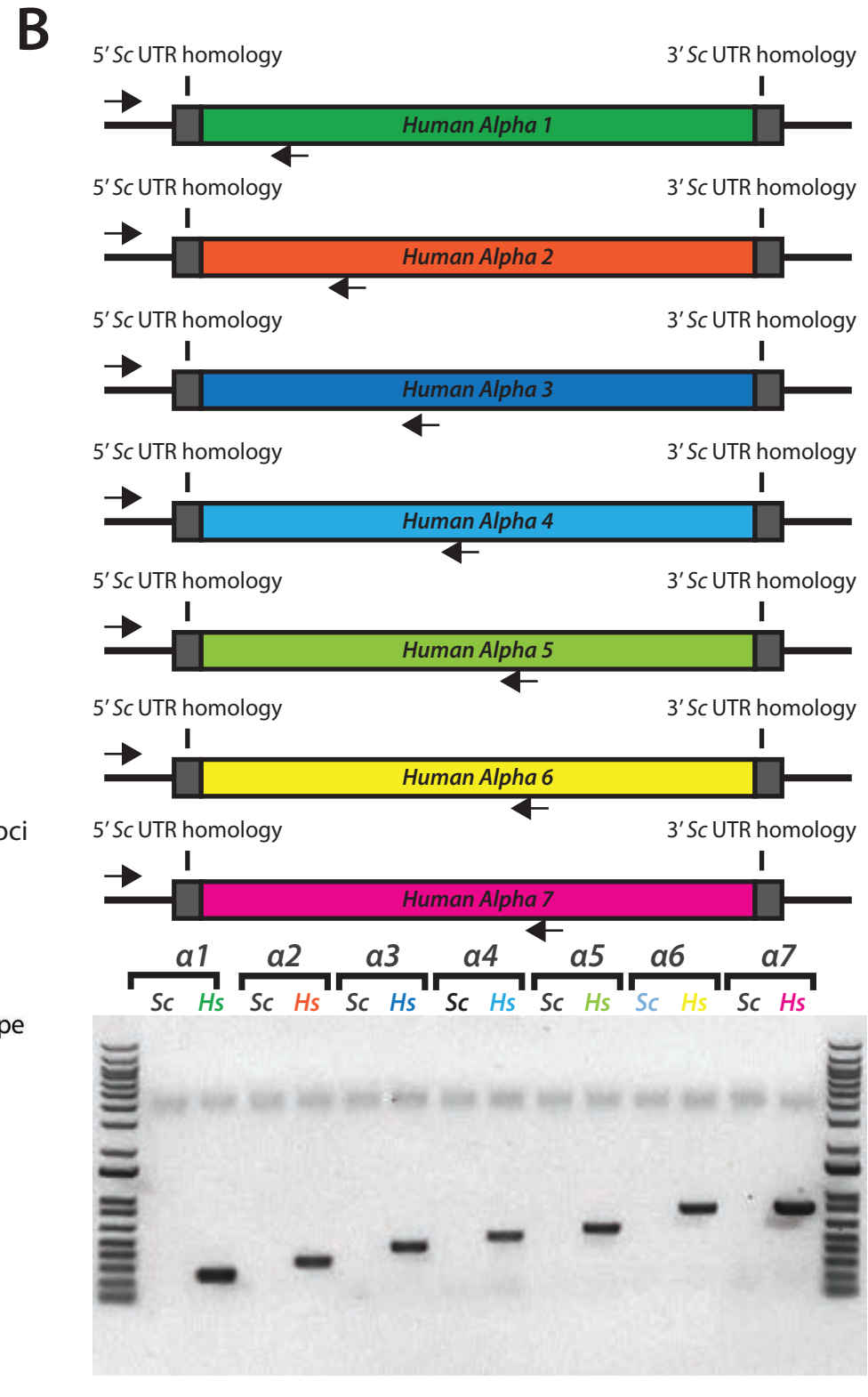
Figure S2

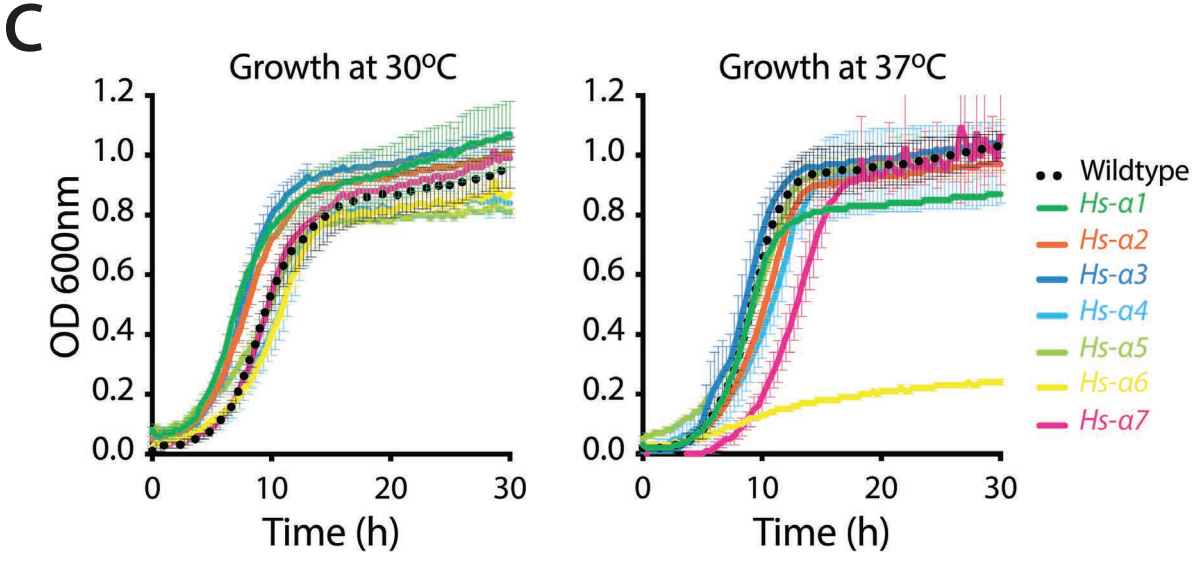

Homozygous loci
Heterozygous loci
Wildtype locus
1
1
1
1
Engineered locus
1
i
i
HDR Efficiency?
NO CRISPR
pCas9-gRNA¹
NO CRISPR
pCas9-gRNA¹
Gene Drive
Normal inheritance
Lethality
Normal inheritance
50% WT: 50% Eng. loci
100% conversion to Engineered loci?

555 **Figure S2. MERGE enables the highly efficient conversion of heterozygous to**
556 **homozygous loci.**
557 Every locus in yeast follows a normal (1:1) inheritance pattern of segregation. However, CRISPR
558 reagent targeted to a haploid or a diploid locus leads to lethality. Similarly, a heterozygous yeast
559 locus follows 1 (1, wild-type) : 1 (i, engineered) inheritance pattern. CRISPR reagent targeted to
560 one of the heterozygous alleles initiates homologous recombination using the CRISPR-resistance
561 homologous chromosome as a repair template, thus enabling efficient conversion to homozygous
562 loci, mimicking a gene drive. The subsequent segregation pattern of 4 : 0 allows propagation of
563 only desired loci.

Figure S3

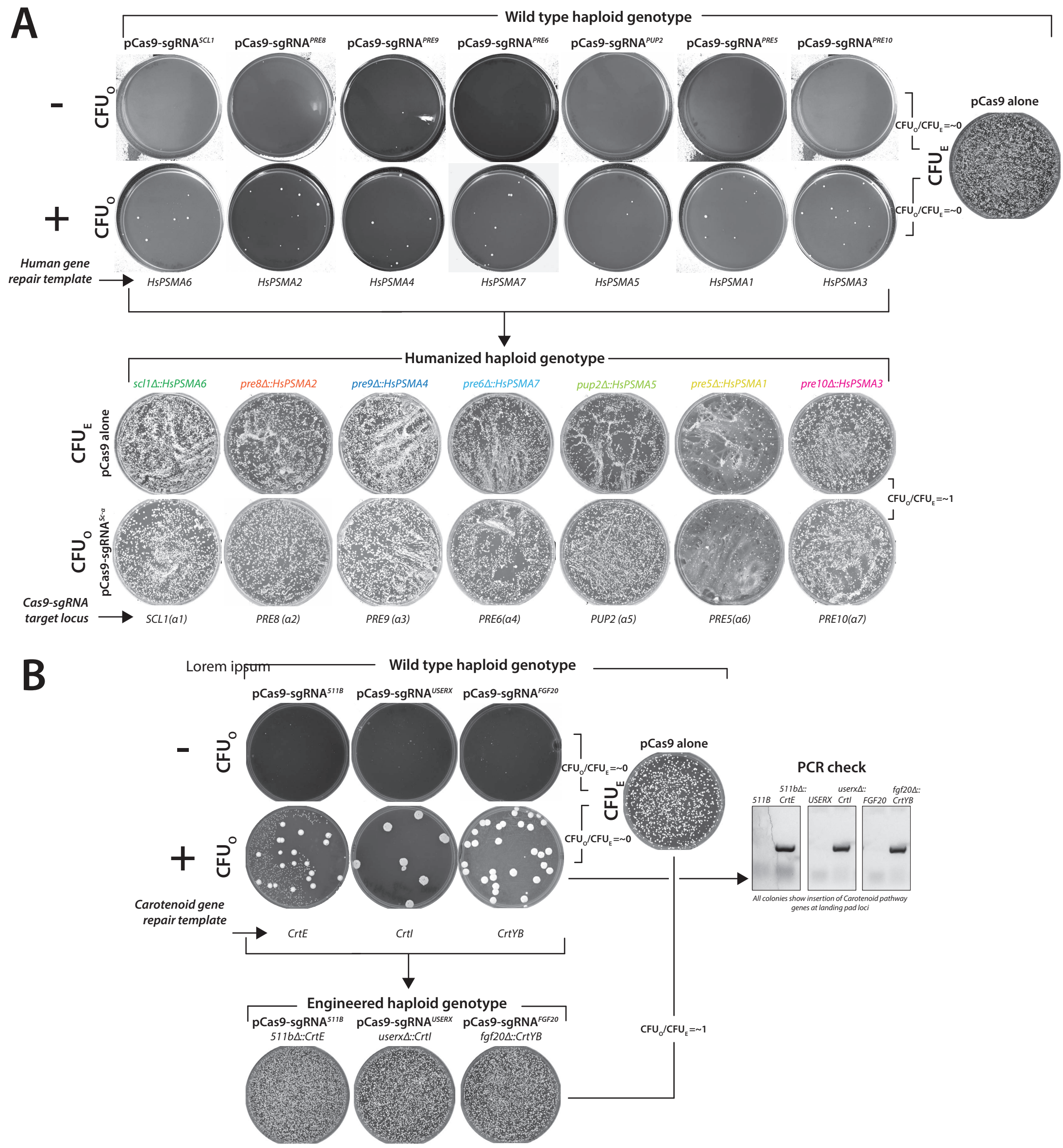

A
CRISPR-Cas9-sgRNA WT locus +
Human gene repair template
HsPSMA6
HsPSMA2
HsPSMA4
HsPSMA7
HsPSMA5
HsPSMA1
HsPSMA3
pCas9-sgRNA plasmid
WT yeast
Hsα1 PSMA6
Hsα2 PSMA2
Hsα3 PSMA4
Hsα4 PSMA7
Hsα5 PSMA5
Hsα6 PSMA1
Hsα7 PSMA3
Human gene repair template
CRISPR reagent
Humanized yeast strains with human α proteasome genes at their native yeast loci


B
5′ Sc UTR homology
3′ Sc UTR homology
Human Alpha 1
Human Alpha 2
Human Alpha 3
Human Alpha 4
Human Alpha 5
Human Alpha 6
Human Alpha 7
α1 α2 α3 α4 α5 α6 α7
Sc Hs Sc Hs Sc Hs Sc Hs Sc Hs Sc Hs Sc Hs


C
Growth at 30ºC
Growth at 37ºC
OD 600nm
Time (h)
Wildtype
Hs-α1
Hs-α2
Hs-α3
Hs-α4
Hs-α5
Hs-α6
Hs-α7

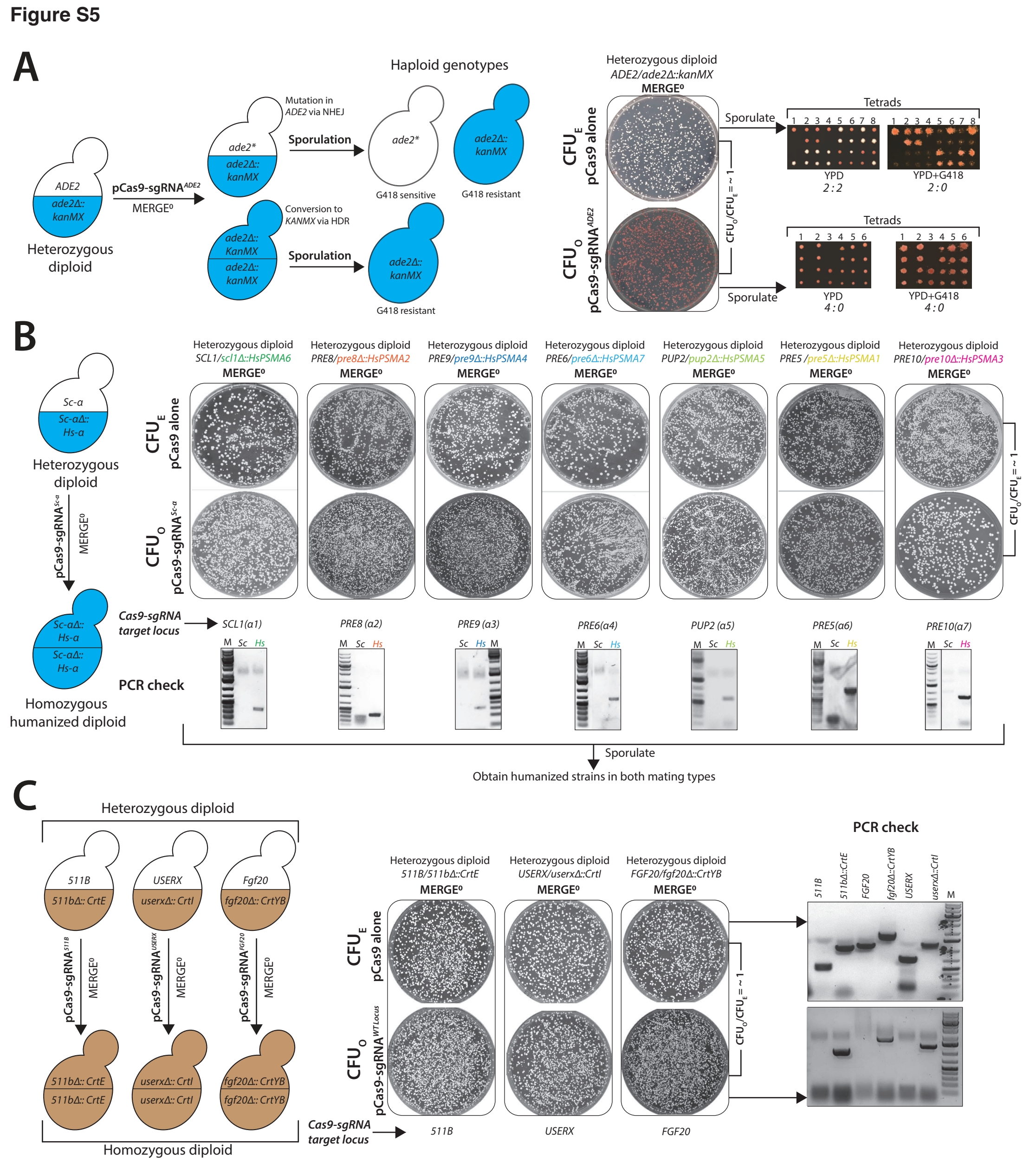
564 **Figure S3. General scheme of CRISPR-Cas9-based gene replacement.**
565 **(A)** For each gene, a 'knockout' (KO) plasmid (CRISPR reagent) is constructed that expresses
566 the Cas9 endonuclease and a sgRNA with a targeting sequence to the gene of interest. To replace
567 the gene of interest, the CRISPR reagent is co-transformed into yeast with a linear repair DNA,
568 consisting of the replacing gene sequence flanked by long (60-200bp) homology sequences to
569 the flanking region of the genome. Cas9 and the sgRNA are co-expressed and target the gene of
570 interest for cutting, allowing repair by homologous recombination (HR) with the provided repair
571 DNA. **(B)** After obtaining transformants, sequencing across the junction reveals successful
572 replacement (as shown for the human α proteasome subunits replaced at the native yeast loci).
573 PCR for the yeast locus (*Sc*) is negative, whereas for a human gene (*Hs*) is positive. **(C)** The
574 growth of singly-humanized α-subunits is not significantly impaired. Each curve represents an
575 averaging of 4 replicates for each indicated strain, grown in YPD for 24 hours. No strains exhibit
576 significantly impaired growth, though strains with human subunits α4, 5, and 6 show some
577 extension of lag phase at 30°C and marginally slowed doubling times at 30°C. The behavior is
578 similar at 37°C except in *Hs-α6,* which exhibited a significant growth defect.

Figure S4


A
Wild type haploid genotype
pCas9-sgRNA^SCL1
pCas9-sgRNA^PRE8
pCas9-sgRNA^PRE9
pCas9-sgRNA^PRE6
pCas9-sgRNA^PUP2
pCas9-sgRNA^PRE5
pCas9-sgRNA^PRE10
-
+
CFU_O
CFU_O/CFU_E=~0
CFU_E
pCas9 alone
Human gene repair template
HsPSMA6
HsPSMA2
HsPSMA4
HsPSMA7
HsPSMA5
HsPSMA1
HsPSMA3
Humanized haploid genotype
scl1Δ::HsPSMA6
pre8Δ::HsPSMA2
pre9Δ::HsPSMA4
pre6Δ::HsPSMA7
pup2Δ::HsPSMA5
pre5Δ::HsPSMA1
pre10Δ::HsPSMA3
CFU_E pCas9 alone
CFU_O pCas9-sgRNA^Sc-α
CFU_O/CFU_E=~1
Cas9-sgRNA target locus
SCL1(α1)
PRE8 (α2)
PRE9 (α3)
PRE6(α4)
PUP2 (α5)
PRE5(α6)
PRE10(α7)
B
Lorem ipsum
Wild type haploid genotype
pCas9-sgRNA^511B
pCas9-sgRNA^USERX
pCas9-sgRNA^FGF20
-
+
CFU_O
CFU_O/CFU_E=~0
CFU_E
pCas9 alone
PCR check
511B
511bΔ::CrtE
USERX
userxΔ::CrtI
FGF20
fgf20Δ::CrtYB
All colonies show insertion of Carotenoid pathway genes at landing pad loci
Carotenoid gene repair template
CrtE
CrtI
CrtYB
Engineered haploid genotype
pCas9-sgRNA^511B
511bΔ::CrtE
pCas9-sgRNA^USERX
userxΔ::CrtI
pCas9-sgRNA^FGF20
fgf20Δ::CrtYB
CFU_O/CFU_E=~1

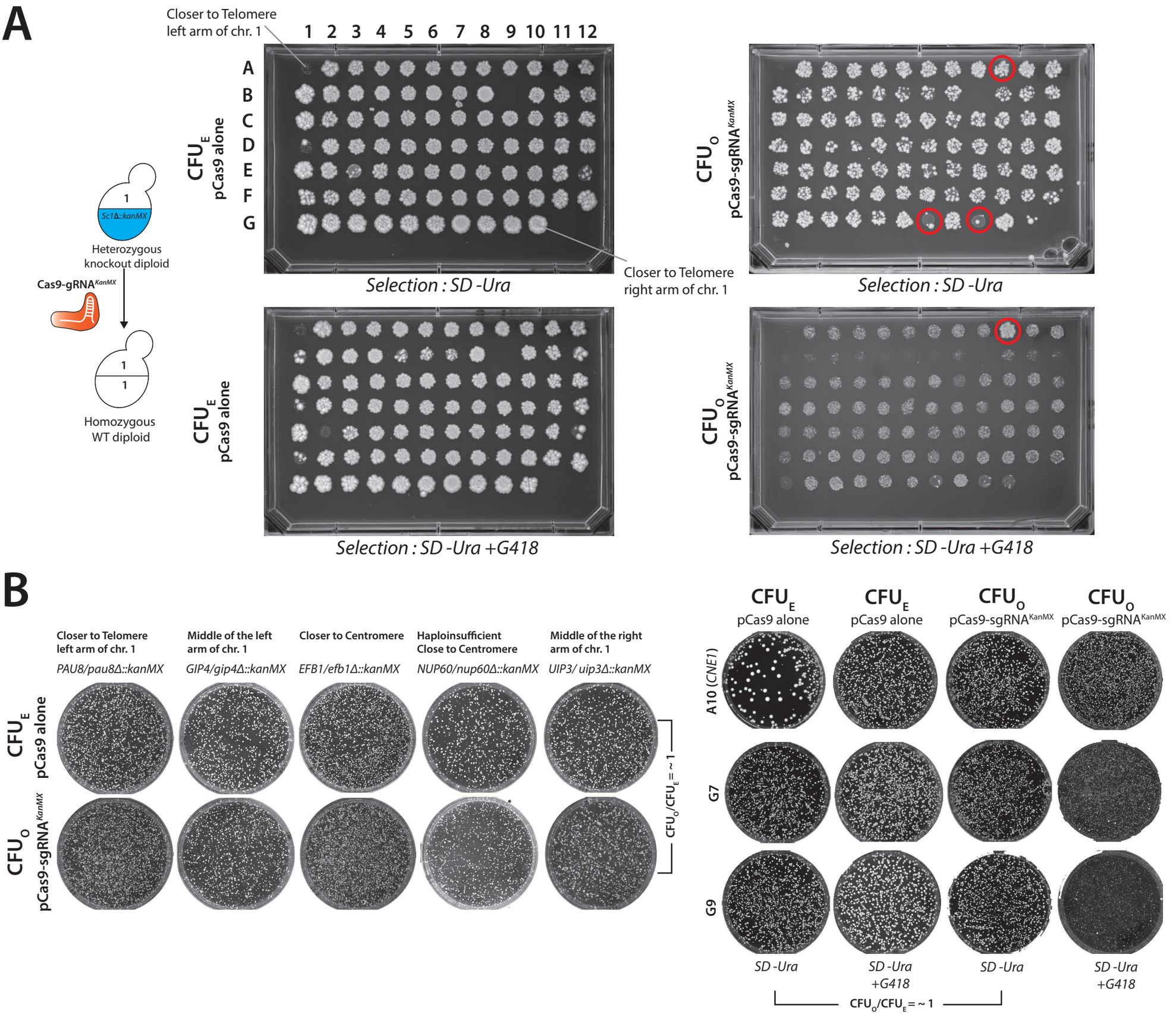
579 **Figure S4. General scheme of CRISPR-Cas9-based gene replacement, CRISPR-mediated**
580 **sensitivity and resistance of yeast loci.**
581 **(A)** CRISPR reagent (**pCas9-sgRNA^{locus}**) targeted to any yeast locus often leads to lethality with
582 few surviving colonies (**CFU_O**). The corresponding transformation of a control plasmid (**pCas9**
583 **alone**) estimates the transformation efficiency (**CFU_E**). CRISPR reagents targeting 7 yeast α
584 proteasome genes show lethality (**CFU_O/CFU_E = ~0**) with or without the repair template. After
585 humanizing the corresponding yeast loci, each single-humanized haploid strain shows resistance
586 to further targeting by the corresponding CRISPR reagent, respectively (**CFU_O/CFU_E = ~1**). **(B)**
587 CRISPR reagents targeted to landing pad loci (*511B*, *USERX* and *FGF20*) enabled editing by
588 introducing carotenoid gene transcription units *CrtE*, *CrtI* and *CrtYB* (provided as PCR fragment
589 repair templates). The engineered strains show resistance to further cutting by the corresponding
590 CRISPR reagents (**CFU_O/CFU_E = ~1**). PCR-based genotyping confirmed the integration of each
591 carotenoid transcription unit for every colony tested.

Figure S5


A
Haploid genotypes
Heterozygous diploid
ADE2/ade2Δ::kanMX
MERGE⁰
ADE2
ade2Δ::kanMX
Heterozygous diploid
pCas9-sgRNA^ADE2
MERGE⁰
Mutation in ADE2 via NHEJ
ade2*
ade2Δ::kanMX
Sporulation
ade2*
ade2Δ::kanMX
G418 sensitive
G418 resistant
Conversion to KANMX via HDR
ade2Δ::KanMX
ade2Δ::kanMX
Sporulation
ade2Δ::kanMX
G418 resistant
CFU_E pCas9 alone
CFU_O pCas9-sgRNA^ADE2
CFU_O/CFU_E = ~ 1
Sporulate
Tetrads
1 2 3 4 5 6 7 8
YPD
2 : 2
YPD+G418
2 : 0
Sporulate
Tetrads
1 2 3 4 5 6
YPD
4 : 0
YPD+G418
4 : 0
B
Heterozygous diploid
Sc-α
Sc-αΔ::Hs-α
pCas9-sgRNA^Sc-α
MERGE⁰
Homozygous humanized diploid
Heterozygous diploid SCL1/scl1Δ::HsPSMA6
Heterozygous diploid PRE8/pre8Δ::HsPSMA2
Heterozygous diploid PRE9/pre9Δ::HsPSMA4
Heterozygous diploid PRE6/pre6Δ::HsPSMA7
Heterozygous diploid PUP2/pup2Δ::HsPSMA5
Heterozygous diploid PRE5 /pre5Δ::HsPSMA1
Heterozygous diploid PRE10/pre10Δ::HsPSMA3
MERGE⁰
CFU_E pCas9 alone
CFU_O pCas9-sgRNA^Sc-α
CFU_O/CFU_E = ~ 1
Cas9-sgRNA target locus
SCL1(α1)
PRE8 (α2)
PRE9 (α3)
PRE6(α4)
PUP2 (α5)
PRE5(α6)
PRE10(α7)
M Sc Hs
PCR check
Sporulate
Obtain humanized strains in both mating types
C
Heterozygous diploid
511B
511bΔ:: CrtE
USERX
userxΔ:: CrtI
Fgf20
fgf20Δ:: CrtYB
pCas9-sgRNA^511B
pCas9-sgRNA^USERX
pCas9-sgRNA^FGF20
MERGE⁰
511bΔ:: CrtE
511bΔ:: CrtE
userxΔ:: CrtI
userxΔ:: CrtI
fgf20Δ:: CrtYB
fgf20Δ:: CrtYB
Homozygous diploid
Heterozygous diploid 511B/511bΔ::CrtE
Heterozygous diploid USERX/userxΔ::CrtI
Heterozygous diploid FGF20/fgf20Δ::CrtYB
MERGE⁰
CFU_E pCas9 alone
CFU_O pCas9-sgRNA^WT Locus
CFU_O/CFU_E = ~ 1
Cas9-sgRNA target locus
511B
USERX
FGF20
PCR check
511B
511bΔ::CrtE
FGF20
fgf20Δ::CrtYB
USERX
userxΔ::CrtI
M

592 **Figure S5. MERGE^0 enables the efficient conversion of single heterozygous to**
593 **homozygous loci.**
594 **(A)** A schematic of possible genotypes after the *ADE2* locus in a heterozygous diploid
595 *ADE2*/*ade2Δ::kanMX* strain is subjected to a CRISPR-Cas9-mediated DSB followed by
596 sporulation. The repair via NHEJ could mutate the locus (*ade2**) or repair via HDR converts the
597 locus to *ade2Δ::kanMX*, either scenario leads to a red colony phenotype. However, post
598 sporulation, the prior scenario will show only 50% G418 resistant colonies compared to the latter
599 outcome, which will exhibit 100% G418 resistance. Tetrad dissection of pooled cells from
600 heterozygous diploid strains before **MERGE^0** (**pCas9 alone**) or after **MERGE^0** (**pCas9-**
601 **sgRNA^{ADE2}**) shows that before **MERGE^0** cells appear as 2 : 2 (red : white) and 2 : 2 (G418-
602 resistant : G418-sensitive) phenotype. However, after **MERGE^0,** cells show 4 : 0 (red : white)
603 phenotype. Furthermore, all 4 spores in the post **MERGE^0** strain show 4 : 0 G418 resistance. **(B)**
604 Schematic shows CRISPR reagent targeted to one of the heterozygous alleles (wild type yeast
605 gene) in a diploid strain permits efficient conversion to homozygous humanized loci. **MERGE^0**
606 efficiently converts wild-type yeast loci α proteasome genes to humanized loci (**CFU_O/CFU_E= ~0**).
607 PCR-based genotyping verified the conversion to the CRISPR-resistant humanized loci for every
608 colony tested. **(C)** Schematic shows how CRISPR reagent targeted to the landing pad loci in a
609 heterozygous diploid enables conversion to engineered carotenoid loci. **MERGE^0** converts
610 heterozygous landing pad loci (*511B, USERX,* and *FGF20*) to engineered carotenoid loci
611 (**CFU_O/CFU_E = ~1)**. PCR-based genotyping shows heterozygous status of the loci before
612 **MERGE^0** and conversion to engineered carotenoid loci after **MERGE^0**. Sporulation yields haploids
613 of both mating types with only engineered loci.

Figure S6


A
Closer to Telomere
left arm of chr. 1
1 2 3 4 5 6 7 8 9 10 11 12
A
B
C
D
E
F
G
CFU_E
pCas9 alone
1
Sc1Δ::kanMX
Heterozygous
knockout diploid
Cas9-gRNA^KanMX
1
1
Homozygous
WT diploid
Selection : SD -Ura
Closer to Telomere
right arm of chr. 1
CFU_O
pCas9-sgRNA^KanMX
Selection : SD -Ura
CFU_E
pCas9 alone
Selection : SD -Ura +G418
CFU_O
pCas9-sgRNA^KanMX
Selection : SD -Ura +G418
B
Closer to Telomere
left arm of chr. 1
PAU8/pau8Δ::kanMX
Middle of the left
arm of chr. 1
GIP4/gip4Δ::kanMX
Closer to Centromere
EFB1/efb1Δ::kanMX
Haploinsufficient
Close to Centromere
NUP60/nup60Δ::kanMX
Middle of the right
arm of chr. 1
UIP3/ uip3Δ::kanMX
CFU_E
pCas9 alone
CFU_O
pCas9-sgRNA^KanMX
CFU_O/CFU_E = ~ 1
CFU_E
pCas9 alone
CFU_E
pCas9 alone
CFU_O
pCas9-sgRNA^KanMX
CFU_O
pCas9-sgRNA^KanMX
A10 (CNE1)
G7
G9
SD -Ura
SD -Ura
+G418
SD -Ura
SD -Ura
+G418
CFU_O/CFU_E = ~ 1

614 **Figure S6. MERGE^0 is efficient at converting many yeast heterozygous to homozygous**
615 **loci independent of the position on Chromosome 1.**
616 **(A)** Schematic shows a CRISPR reagent (**pCas9-sgRNA^{KanMX}**) targeted to the KanMX allele in a
617 hetKO strain converts the locus to homozygous wild-type. The hetKO strains were arrayed as
618 genes close to the left-arm of a telomere through the centromere to the right arm of a telomere.
619 The strains were transformed with **pCas9 alone** (**CFU_E**) or with **pCas9-sgRNA^{KanMX}** (**CFU_O**) and
620 spotted on either SD-Ura or SD-Ura with G418 selection. Every genotype shows spots with a
621 similar density of colonies showing no lethality due to the CRISPR reagent (**CFU_O/CFU_E = ~1**).
622 The spots on SD-Ura with G418 selection show G418 resistance in the case of **pCas9 alone**
623 transformants. However, all but one (A10 - *CNE1*) strain transformed with **pCas9-sgRNA^{KanMX}**
624 lost the *KanMX* cassette resulting in G418 sensitivity. **(B) MERGE^0** mediated conversion of loci
625 was tested on several unique hetKO genotypes on Petri plates, demonstrating agreement with
626 96-well spot assays. All tested strains show no lethality due to the CRISPR reagent (**CFU_O/CFU_E**
627 **= ~1**). Ambiguous spots on a 96-well plate (A10, G7 and G9) were further analyzed on Petri
628 plates. G7 and G9 spots show no lethality due to CRISPR reagent (**CFU_O/CFU_E = ~1**) and loss of
629 *KanMX* cassette. While the A10 (*CNE1*) strain is resistant to the CRISPR reagent, the cells retain
630 G418 resistance. It is unclear why this behavior occurs.

Figure S7

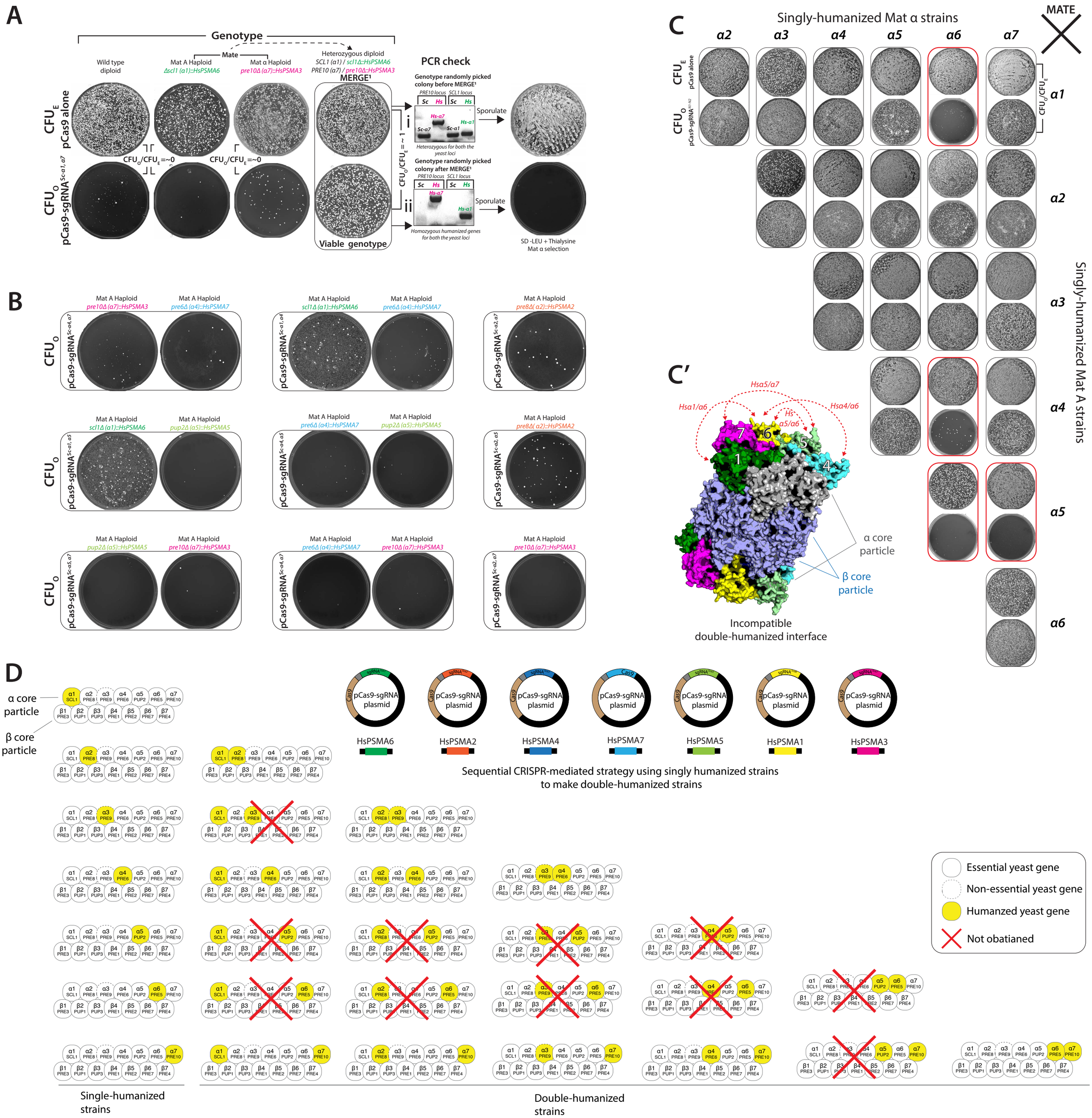
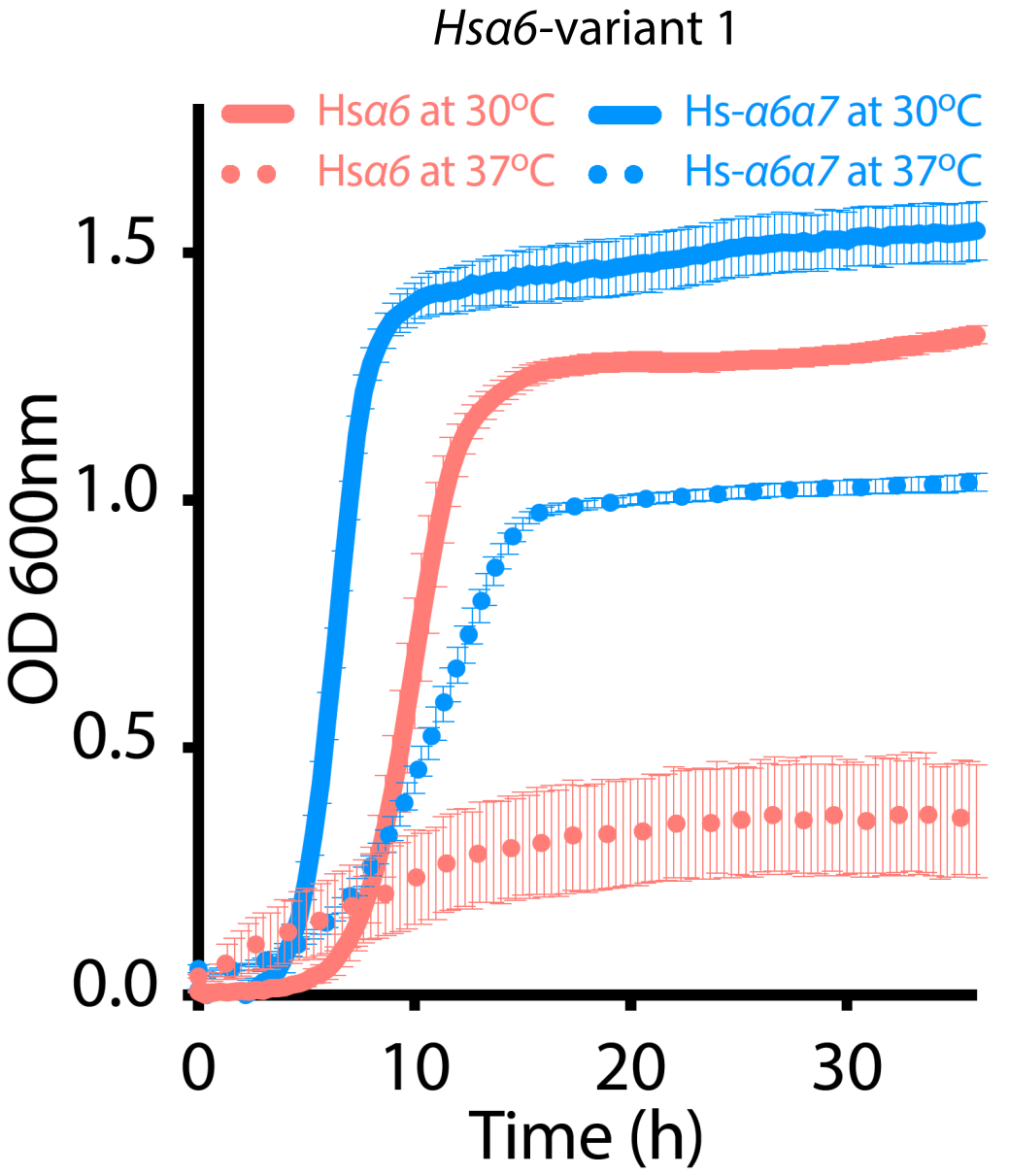

A
Genotype
Mate
Wild type diploid
Mat A Haploid
Δscl1 (a1)::HsPSMA6
Mat α Haploid
pre10Δ (a7)::HsPSMA3
Heterozygous diploid
SCL1 (a1) / scl1Δ::HsPSMA6
PRE10 (a7) / pre10Δ::HsPSMA3
MERGE¹
CFU_E pCas9 alone
CFU_O pCas9-sgRNA^Sc-a1, a7
CFU_O/CFU_E=~0
CFU_O/CFU_E=~0
Viable genotype
CFU_O/CFU_E = ~ 1
PCR check
Genotype randomly picked colony before MERGE¹
PRE10 locus
SCL1 locus
Sc Hs Sc Hs
Hs-a7
Sc-a7
Sc-a1
Hs-a1
Heterozygous for both the yeast loci
Sporulate
Genotype randomly picked colony after MERGE¹
Homozygous humanized genes for both the yeast loci
SD -LEU + Thialysine
Mat α selection
B
Mat A Haploid
pre10Δ (a7)::HsPSMA3
pre6Δ (a4)::HsPSMA7
scl1Δ (a1)::HsPSMA6
pre8Δ(a2)::HsPSMA2
pup2Δ (a5)::HsPSMA5
CFU_O
pCas9-sgRNA^Sc-a4, a7
pCas9-sgRNA^Sc-a1, a4
pCas9-sgRNA^Sc-a2, a7
pCas9-sgRNA^Sc-a1, a5
pCas9-sgRNA^Sc-a4, a5
pCas9-sgRNA^Sc-a2, a5
pCas9-sgRNA^Sc-a5, a7
C
Singly-humanized Mat α strains
MATE
a2
a3
a4
a5
a6
a7
CFU_E pCas9 alone
CFU_O pCas9-sgRNA^Y1/N2
CFU_O/CFU_E
a1
Singly-humanized Mat A strains
C′
Hsa5/a7
Hsa1/a6
Hsa4/a6
Hs a5/a6
α core particle
β core particle
Incompatible double-humanized interface
D
α core particle
β core particle
pCas9-sgRNA plasmid
HsPSMA6
HsPSMA2
HsPSMA4
HsPSMA7
HsPSMA5
HsPSMA1
HsPSMA3
Sequential CRISPR-mediated strategy using singly humanized strains to make double-humanized strains
Essential yeast gene
Non-essential yeast gene
Humanzed yeast gene
Not obatianed
Single-humanized strains
Double-humanized strains

Figure S7. MERGE^0 is efficient at simultaneously combining two heterozygous to homozygous loci.

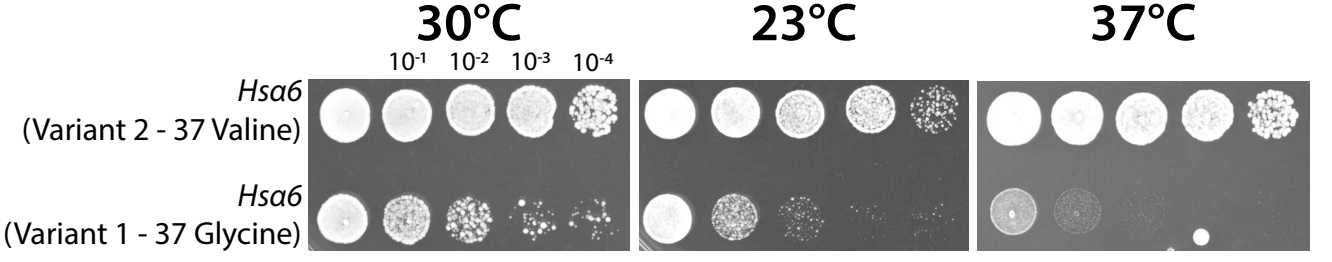
(A) Double-sgRNA CRISPR reagent (**pCas9-sgRNA$^{Sc\alpha1,\alpha7}$**) targeted to two wild type yeast loci (proteasome *α*1 and *α*7 genes) quantifies the efficiency of **MERGE^1**. The transformation of **pCas9-sgRNA$^{Sc\alpha1,\alpha7}$** in the wild type, and single-humanized Hs*α*1 or Hs*α*7 strains is lethal, suggesting ON-target activity (**CFU_O/CFU_E = ~0**). However, the transformation of **pCas9-sgRNA$^{Sc\alpha1,\alpha7}$** in diploid heterozygous humanized strain shows no lethality (**CFU_O/CFU_E = ~1**). PCR-based genotyping of surviving colonies after **MERGE^1** shows conversion of both yeast to the humanized loci compared to before **MERGE^1. (B)** All tested single-humanized strains transformed with double-sgRNA CRISPR reagents show lethality, demonstrating that selection only allows survival of viable paired genotypes. **(C)** Mating each single-humanized strain of opposite mating-types (obtained via **MERGE^0**) provided 21 heterozygous genotypes to test **MERGE^1**. Heterozygous diploid humanized strains transformed with double-sgRNA CRISPR reagents specific to each yeast allele allow testing viability of all paired-humanized genotypes (**CFU_O**). If the genotype is permitted, the strategy enables survival similar to the pCas9 alone transformation (**CFU_E**). While the majority of combined double-humanized genotypes are viable (**CFU_O/CFU_E = ~1**), specific genotypes (*Hsα1α6*, *Hsα4α6*, *Hsα5α6* and *Hsα5α7*) were inviable, suggesting incompatible genotypes. Other combinations, such as those involving the genotypes with Hs*α*2, showed a slow-growth phenotype. **(C')** Yeast proteasome structure (PDB-1RYP) with *α* and β cores show the humanized *α* subunits that are incompatible as paired genotypes. Except for *α5* and *α6*, the subunits are missing neighboring interactions within the α core. **(D)** A sequential strategy using single-humanized strains acquired several pairwise combinations of human α subunits. However, several other genotypes could not be obtained, revealing the drawback of the method. **MERGE^1** is far more efficient in identifying incompatible paired genotypes.

Figure S8

A

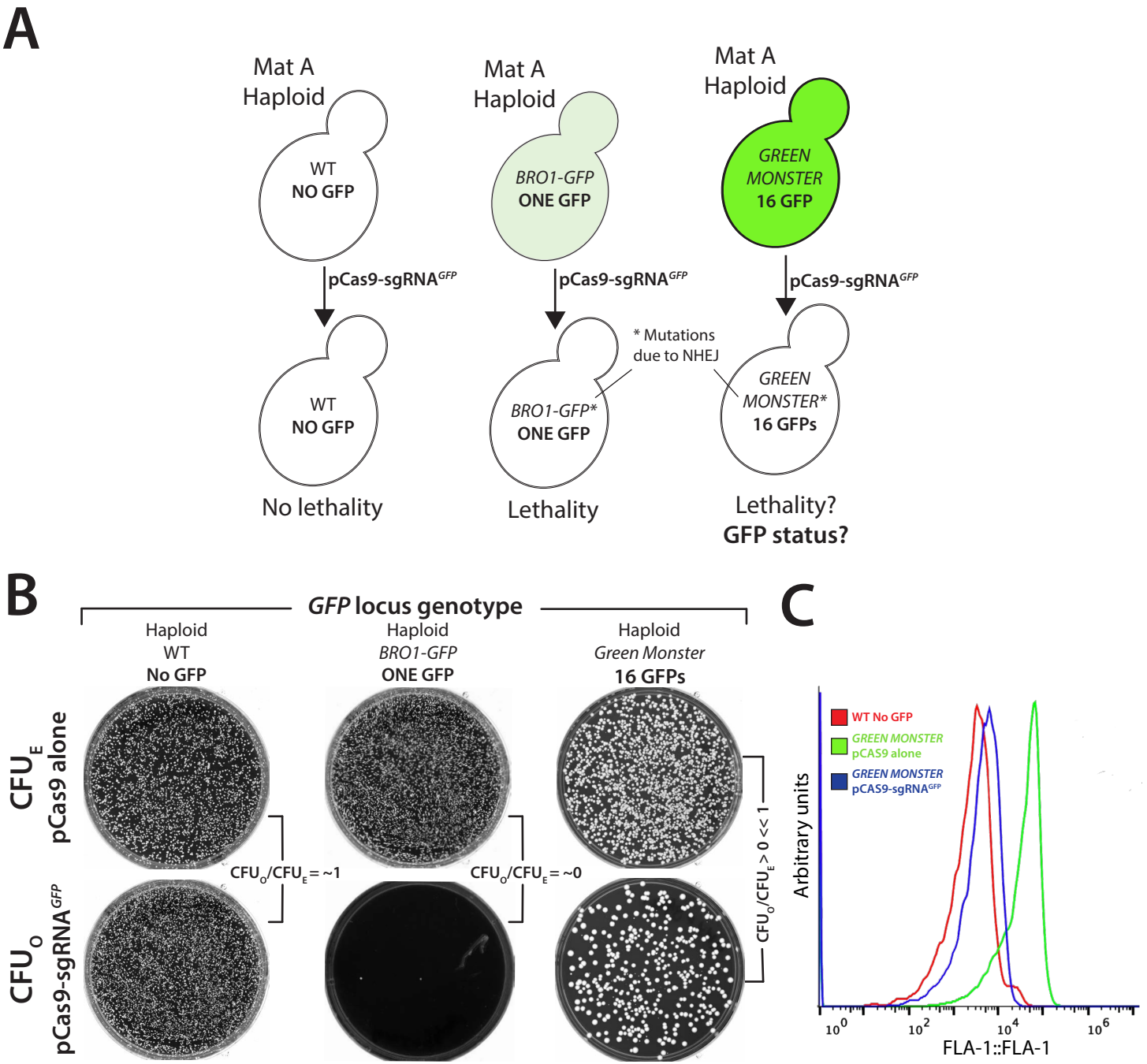

Hsa6-variant 1
Hs*a6* at 30°C
Hs-*a6a7* at 30°C
Hs*a6* at 37°C
Hs-*a6a7* at 37°C
OD 600nm
1.5
1.0
0.5
0.0
0
10
20
30
Time (h)

B


30°C
23°C
37°C
10^{-1} 10^{-2} 10^{-3} 10^{-4}
Hsa6
(Variant 2 - 37 Valine)
Hsa6
(Variant 1 - 37 Glycine)

655 **Figure S8. The variant-specific phenotype of single-humanized Hs*α6* (*PSMA1*) yeast**
656 **strain.**
657 **(A)** A growth assay performed on a single-humanized *Hsα6* strain (Variant 1, amino acid residue
658 37>Glycine) shows a temperature-sensitive (TS) phenotype at 37°C (dotted red line) compared
659 to growth at 30°C (solid red line). The combined genotype of Hs*α*6,*α*7 partially rescues the TS
660 phenotype at 37°C (dotted blue line) compared to the growth at 30°C (solid blue line). **(B)**
661 Spotted dilutions of single-humanized *Hsα6* strains (Variant 1, amino acid residue 37>Glycine)
662 and (Variant 2, amino acid residue 37>Valine) show Variant 2 with no growth defect at 30°C,
663 23°C and 37°C compared to Variant 1.

Figure S9


A
Mat A
Haploid
WT
NO GFP
Mat A
Haploid
BRO1-GFP
ONE GFP
Mat A
Haploid
GREEN
MONSTER
16 GFP
pCas9-sgRNA^GFP
pCas9-sgRNA^GFP
pCas9-sgRNA^GFP
* Mutations
due to NHEJ
WT
NO GFP
BRO1-GFP*
ONE GFP
GREEN
MONSTER*
16 GFPs
No lethality
Lethality
Lethality?
GFP status?
B
GFP locus genotype
Haploid
WT
No GFP
Haploid
BRO1-GFP
ONE GFP
Haploid
Green Monster
16 GFPs
CFU_E
pCas9 alone
CFU_O
pCas9-sgRNA^GFP
CFU_O/CFU_E = ~1
CFU_O/CFU_E = ~0
CFU_O/CFU_E > 0 << 1
C
WT No GFP
GREEN MONSTER
pCAS9 alone
GREEN MONSTER
pCAS9-sgRNA^GFP
Arbitrary units
10^0
10^2
10^4
10^6
FLA-1::FLA-1

664 **Figure S9. MERGE is scalable to target multiple yeast loci simultaneously.**
665 **(A)** The schematic shows a CRISPR-based strategy targeting one or many genetic loci in yeast.
666 Transformation of a CRISPR reagent (**pCas9-sgRNA^{GFP}**) targeting a GFP cassette in a Green
667 Monster strain tests the scalability. Transformation in the wild type (no GFP) strain should show
668 no lethality, whereas strains harboring single or multiple GFPs should show a lethal phenotype.
669 **(B)** The **pCas9-sgRNA^{GFP}** transformation in wild type yeast showed viable cells (**CFU_O/CFU_E**
670 **=~1**). In contrast, transformation in a strain harboring a single GFP cassette was lethal. While the
671 Green Monster strain (GFP monster, 16 GFP cassettes) showed more viable cells than the
672 single-GFP strain upon transformation with **pCas9-sgRNA^{GFP}**. **(C)** Using flow cytometry, the
673 pooled mixture of yeast cells that survive **pCas9-sgRNA^{GFP}** showed significantly reduced GFP
674 expression compared to the pCas9 alone transformed cells, suggesting that nearly every GFP
675 cassette was successfully targeted and mutated via NHEJ.

Figure S10

A

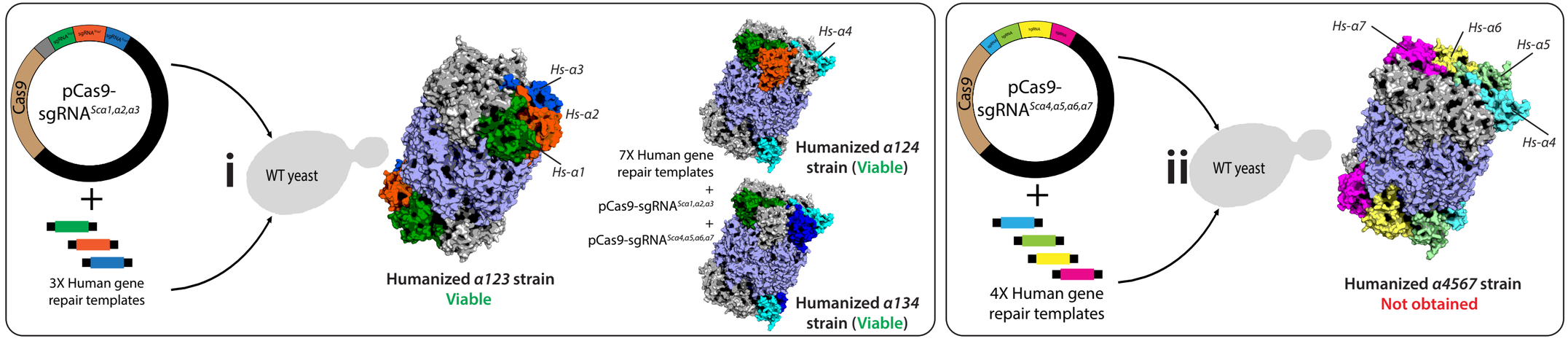
Simultaneous multiple Cas9-sgRNAs & PCR fragment repair templates

↓

Select using CRISPR/Cas9 selection

↓

Genotype surviving colonies



B

Sequential strategy

Humanized *α123* strain

+pCas9-sgRNA^Sca4
+ Hs*α4* repair template

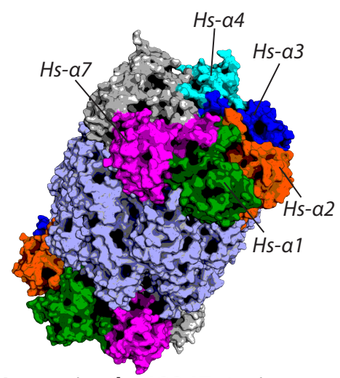
↓

Humanized *α1234* strain

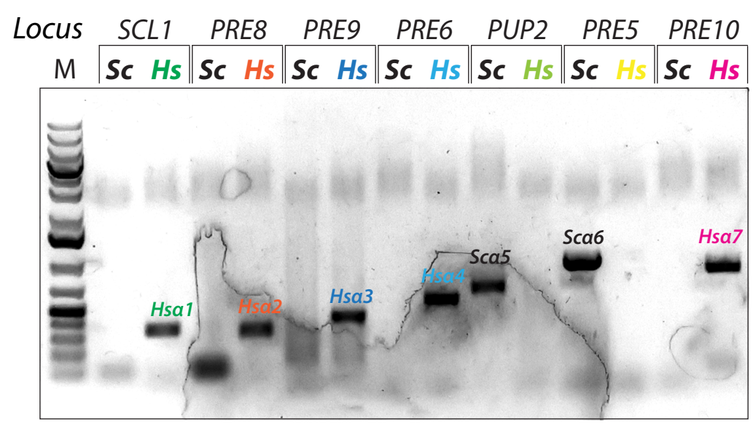
+pCas9-sgRNA^Sca7
+ Hs*α7* repair template

↓

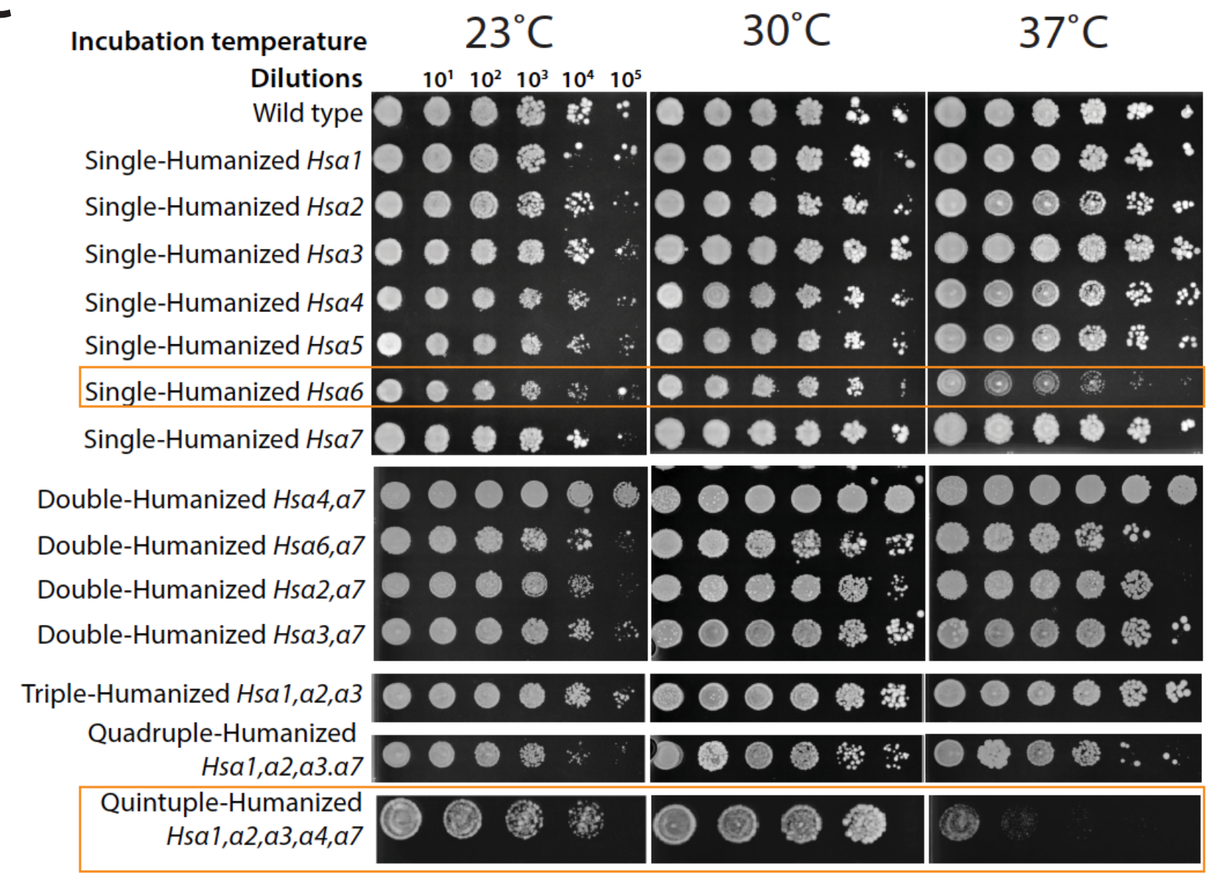
Humanized *α12347* strain



Humanized *α12347* strain

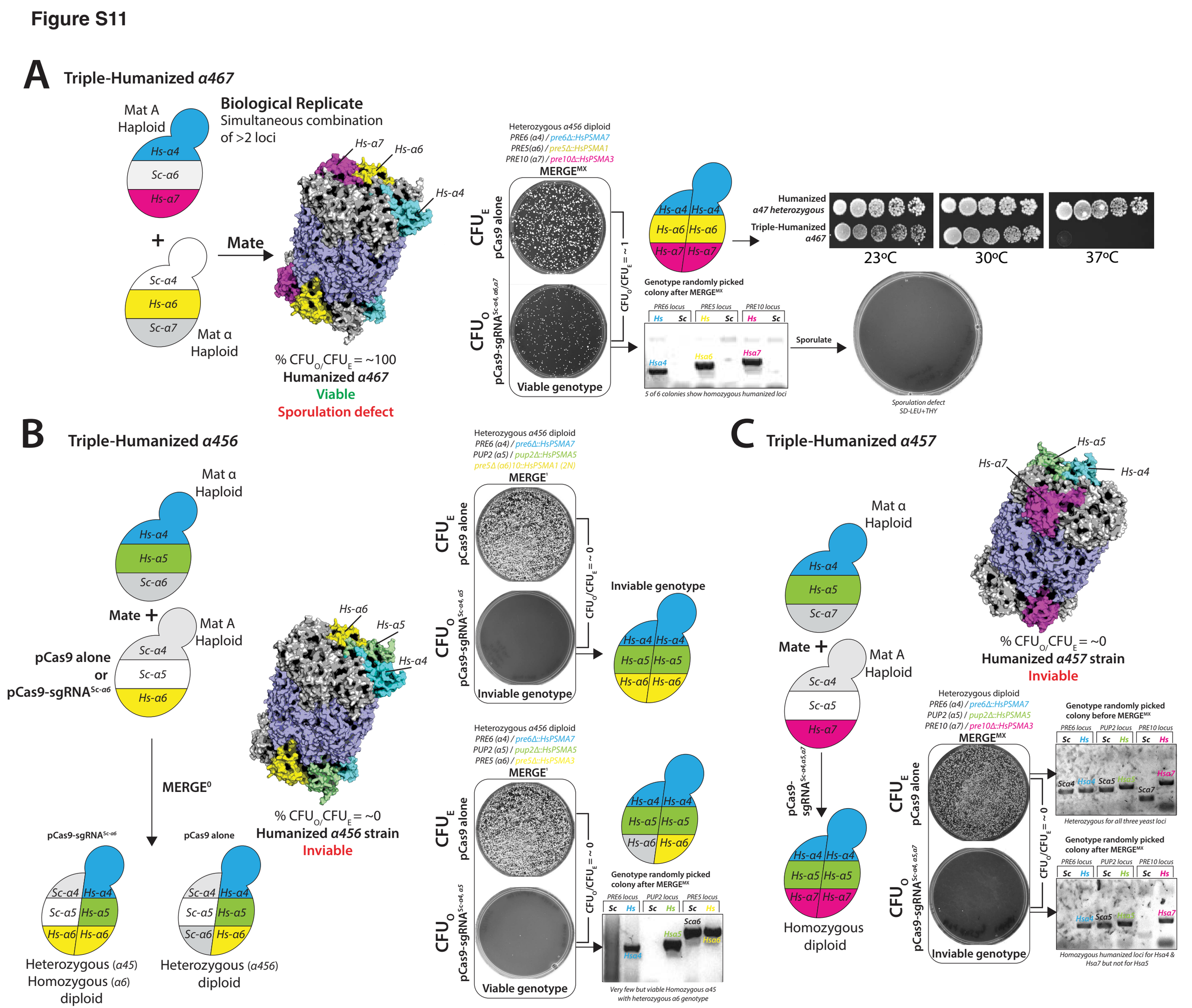


C



676 **Figure S10. Scheme for obtaining triple-, quadruple-, quintuple- and sextuple-humanized**
677 **α-subunit strains *via* a sequential strategy.**
678 **(A)** Two multiple-sgRNA CRISPR plasmids were constructed, one expressing sgRNAs targeting
679 yeast α subunits 1, 2, and 3 (**pCas9-sgRNA**$^{Sc\alpha1,\alpha2,\alpha3}$), and the other expressing sgRNAs targeting
680 α subunits 4, 5, 6, and 7 (**pCas9-sgRNA**$^{Sc\alpha4,\alpha5,\alpha6,\alpha7}$). They were transformed individually with the
681 respective human gene repair DNA. A single colony with triple-humanized *Hs-α1α2,α3* (**i**) was
682 obtained. However, no clone for a quadruple humanized *Hs-α4α5,α6,α7* (**ii**) could be obtained. A
683 similar strategy also generated triple-humanized *Hs-α1α2,α4* and *Hs-α1α3,α4* genotypes. **(B)**
684 Sequential strategy generated a quintuple-humanized *Hs-α1α2,α3*,*α4,α7* strain confirmed via
685 locus-specific PCR showing the presence of human and the absence of yeast loci. Sanger
686 sequencing and whole-genome sequencing confirmed the genotype. **(C)** Spotted dilutions of
687 double-, triple-, quadruple-, quintuple- and sextuple humanized *α* proteasome strains at 23°C,
688 30°C and 37°C show singe-humanized Hs-*α6* (variant 1) and quintuple-humanized *Hs-*
689 *α1,α2,α3*,*α4,α7* strain with a temperature-sensitive growth phenotype.

Figure S11


A
Triple-Humanized α467
Mat A Haploid
Hs-α4
Sc-α6
Hs-α7
+
Sc-α4
Hs-α6
Sc-α7
Mat α Haploid
Mate
Biological Replicate
Simultaneous combination of >2 loci
Hs-α7
Hs-α6
Hs-α4
% CFUO/CFUE = ~100
Humanized α467
Viable
Sporulation defect
Heterozygous α456 diploid
PRE6 (α4) / pre6Δ::HsPSMA7
PRE5(α6) / pre5Δ::HsPSMA1
PRE10 (α7) / pre10Δ::HsPSMA3
MERGEMX
CFUE pCas9 alone
CFUO pCas9-sgRNASc-α4, α6,α7
Viable genotype
CFUO/CFUE = ~ 1
Genotype randomly picked colony after MERGEMX
PRE6 locus
PRE5 locus
PRE10 locus
Hs
Sc
Hsα4
Hsα6
Hsα7
5 of 6 colonies show homozygous humanized loci
Sporulate
Sporulation defect
SD-LEU+THY
Humanized α47 heterozygous
Triple-Humanized α467
23ºC
30ºC
37ºC
B
Triple-Humanized α456
Mat α Haploid
Hs-α4
Hs-α5
Sc-α6
Mate +
Mat A Haploid
Sc-α4
Sc-α5
Hs-α6
pCas9 alone
or
pCas9-sgRNASc-α6
MERGE0
pCas9-sgRNASc-α6
pCas9 alone
Heterozygous (α45)
Homozygous (α6)
diploid
Heterozygous (α456)
diploid
Hs-α6
Hs-α5
Hs-α4
% CFUO/CFUE = ~0
Humanized α456 strain
Inviable
Heterozygous α456 diploid
PRE6 (α4) / pre6Δ::HsPSMA7
PUP2 (α5) / pup2Δ::HsPSMA5
pre5Δ (α6)10::HsPSMA1 (2N)
MERGE1
CFUE pCas9 alone
CFUO pCas9-sgRNASc-α4, α5
Inviable genotype
CFUO/CFUE = ~ 0
Inviable genotype
Heterozygous α456 diploid
PRE6 (α4) / pre6Δ::HsPSMA7
PUP2 (α5) / pup2Δ::HsPSMA5
PRE5 (α6) / pre5Δ::HsPSMA3
MERGE1
CFUE pCas9 alone
CFUO pCas9-sgRNASc-α4, α5
Viable genotype
CFUO/CFUE = ~ 0
Genotype randomly picked colony after MERGEMX
PRE6 locus
PUP2 locus
PRE5 locus
Sc
Hs
Hsα4
Hsα5
Sca6
Hsα6
Very few but viable Homozygous α45 with heterozygous α6 genotype
C
Triple-Humanized α457
Hs-α5
Hs-α7
Hs-α4
% CFUO/CFUE = ~0
Humanized α457 strain
Inviable
Mat α Haploid
Hs-α4
Hs-α5
Sc-α7
Mate +
Mat A Haploid
Sc-α4
Sc-α5
Hs-α7
pCas9-sgRNASc-α4,α5,α7
Homozygous diploid
Heterozygous diploid
PRE6 (α4) / pre6Δ::HsPSMA7
PUP2 (α5) / pup2Δ::HsPSMA5
PRE10 (α7) / pre10Δ::HsPSMA3
MERGEMX
CFUE pCas9 alone
CFUO pCas9-sgRNASc-α4, α5,α7
Inviable genotype
CFUO/CFUE = ~ 0
Genotype randomly picked colony before MERGEMX
PRE6 locus
PUP2 locus
PRE10 locus
Sc
Hs
Sca4
Hsα4
Sca5
Hsα5
Sca7
Hsα7
Heterozygous for all three yeast loci
Genotype randomly picked colony after MERGEMX
Hsα4
Sca5
Hsα5
Hsα7
Homozygous humanized loci for Hsa4 & Hsa7 but not for Hsa5

690 **Figure S11. MERGE tests the combination of several triple-humanized yeast α-subunit**
691 **genotypes.**
692 **(A) MERGE^{MX}** (**Mate and multiplex**) strategy combined 3 loci after mating a strain with 2
693 humanized loci (Hs*α*4,*α*7, Mat A) with a strain carrying 1 humanized locus (Hs*α*6, Mat α) (biological
694 replicate of data shown in **Figure 4E**). The transformation of **pCas9-sgRNA$^{Sc\alpha4,\alpha6,\alpha7}$** in a mated mix
695 selected a diploid triple-humanized strain (**CFU_O/CFU_E = ~1**). PCR-based genotyping of several
696 colonies after **MERGE^{MX}** shows conversion of all yeast to the humanized loci (5 of 6). Spotted
697 dilutions of humanized Hs-*α*4,*α*7 and Hs-*α*4,*α*6,*α*7 strains at 23°C, 30°C and 37°C show the triple-
698 humanized strain with a temperature-sensitive phenotype at 37°C. The strain does not yield viable
699 haploid spores observed as no growth on haploid-specific selection media (SD-LEU+Thialysine).
700 (**B**) Double-humanized *Hs-α4α5* (haploid Mat α) and single-humanized *Hs-α6* (haploid Mat A,
701 harboring a **pCas9 alone** or **pCas9-sgRNA$^{Sc\text{-}\alpha6}$** vector) strains were mated, generating two
702 distinct heterozygous diploid strains, one heterozygous for three loci and another for two with
703 homozygous *Hs-α6* locus. Each strain was transformed with double-sgRNA CRISPR plasmid
704 (**pCas9-sgRNA$^{Sc\alpha4,\alpha5}$**) to perform **MERGE^{1}**. The transformation caused lethality in both strains
705 (**CFU_O/CFU_E =~0**). However, a few survivors in the context of heterozygous *Hs-α6* locus showed
706 the conversion of two loci (*Hs-α4* and *Hs-α7*), suggesting an incompatible triple-humanized
707 genotype but viable double-humanized genotype harboring a heterozygous yeast-human *α6*
708 locus. **(C)** Double-humanized haploid *Hs-α4α5* (Mat α) and *Hs-α7* (Mat A) strains were mated.
709 The resulting heterozygous diploid strain was transformed with a triple-sgRNA CRISPR reagent
710 (**pCas9-sgRNA$^{Sc\alpha4,\alpha5,\alpha7}$**). The transformation caused lethality with few survivors (**CFU_O/CFU_E**
711 **=~0**). Genotyping of the survivors showed the conversion of two 2 of 3 loci (*Hs-α4* and *Hs-α7*),
712 suggesting an incompatible triple-humanized genotype.

Figure S12

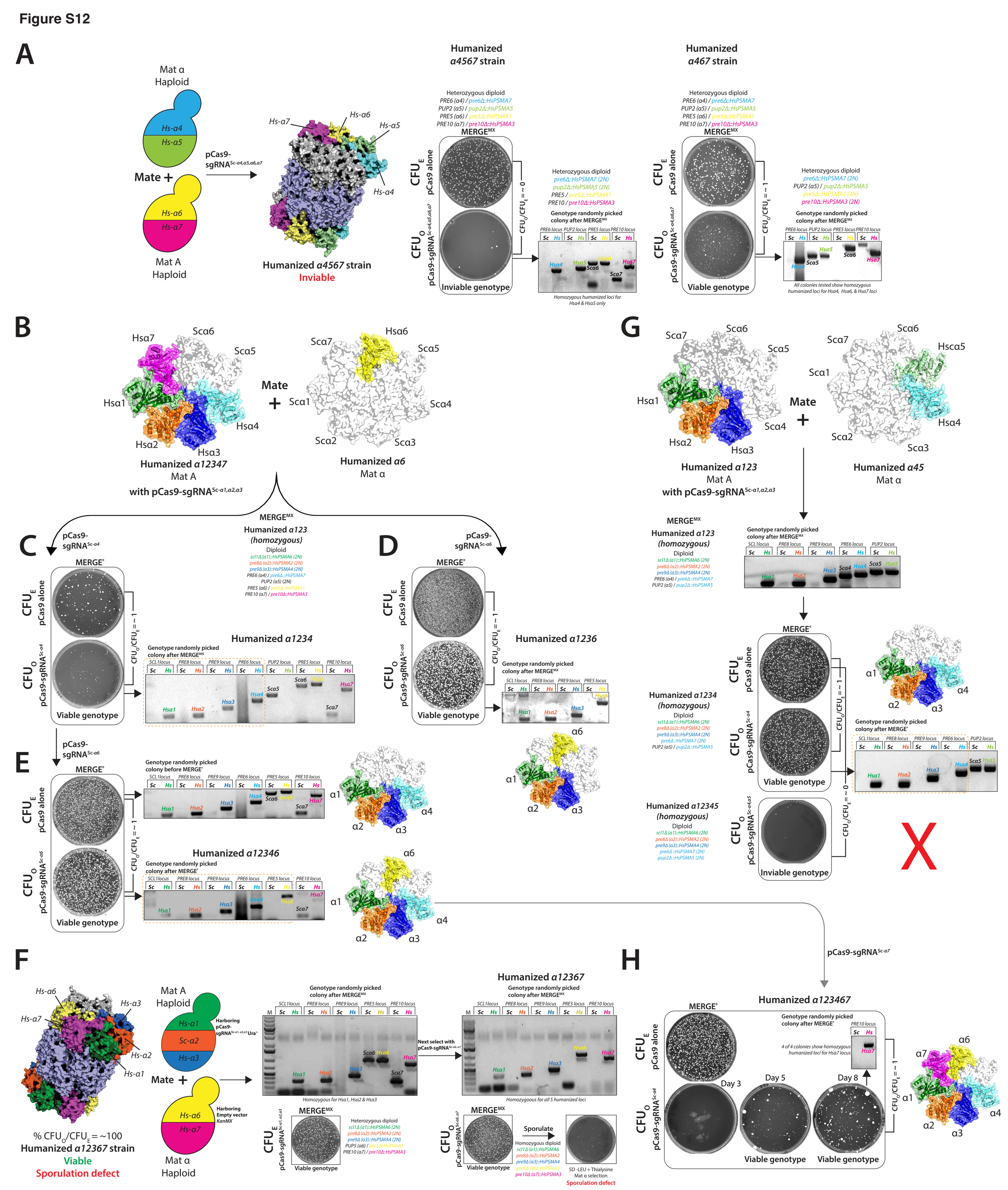
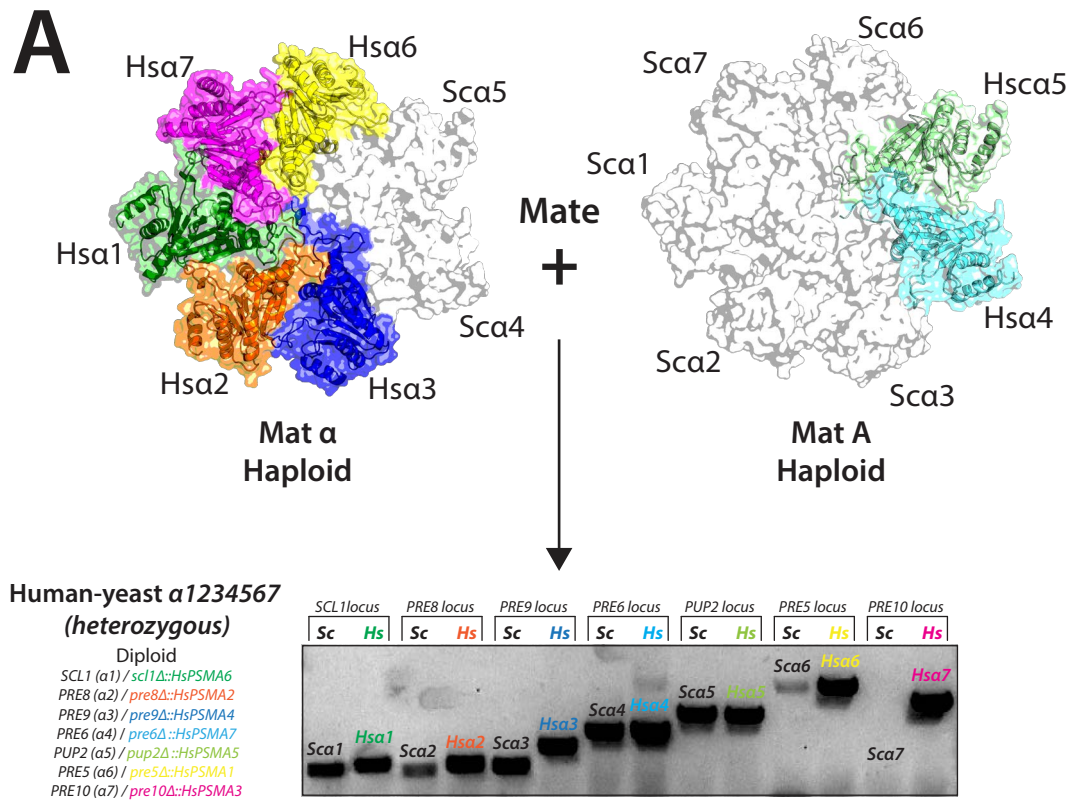

A
Mat α Haploid
Hs-α4
Hs-α5
Mate +
Hs-α6
Hs-α7
Mat A Haploid
pCas9-sgRNA Sc-α4,α5,α6,α7
Humanized α4567 strain
Inviable
Humanized α4567 strain
Humanized α467 strain
Heterozygous diploid
PRE6 (α4) / pre6Δ::HsPSMA7
PUP2 (α5) / pup2Δ::HsPSMA5
PRE5 (α6) / pre5Δ::HsPSMA1
PRE10 (α7) / pre10Δ::HsPSMA3
MERGE MX
CFU E pCas9 alone
CFU O pCas9-sgRNA Sc-α4,α5,α6,α7
CFU O/CFU E = ~ 0
CFU O/CFU E = ~ 1
Inviable genotype
Viable genotype
Heterozygous diploid
pre6Δ::HsPSMA7 (2N)
pup2Δ::HsPSMA5 (2N)
PRE5 / pre5Δ::HsPSMA1
PRE10 / pre10Δ::HsPSMA3
pre5Δ::HsPSMA1 (2N)
pre10Δ::HsPSMA3 (2N)
Genotype randomly picked colony after MERGE MX
Homozygous humanized loci for Hsa4 & Hsa5 only
All colonies tested show homozygous humanized loci for Hsa4, Hsa6, & Hsa7 loci
B
Humanized α12347
Mat A
with pCas9-sgRNA Sc-α1,α2,α3
Mate
Humanized α6
Mat α
C
pCas9-sgRNA Sc-α4
MERGE°
Humanized α123 (homozygous)
Diploid
scl1Δ (α1)::HsPSMA6 (2N)
pre8Δ (α2)::HsPSMA2 (2N)
pre9Δ (α3)::HsPSMA4 (2N)
PRE6 (α4) / pre6Δ::HsPSMA7
PUP2 (α5) (2N)
PRE5 (α6) / pre5Δ::HsPSMA1
PRE10 (α7) / pre10Δ::HsPSMA3
Humanized α1234
Genotype randomly picked colony after MERGE MX
D
pCas9-sgRNA Sc-α6
Humanized α1236
E
pCas9-sgRNA Sc-α6
Genotype randomly picked colony before MERGE°
Humanized α12346
F
Hs-α6
Hs-α3
Hs-α7
Hs-α2
Hs-α1
Mat A Haploid
Harboring pCas9-sgRNA Sc-α1,α2,α3 Ura+
Sc-α2
Mate +
Harboring Empty vector KanMX
Mat α Haploid
% CFU O/CFU E = ~100
Humanized α12367 strain
Viable
Sporulation defect
Genotype randomly picked colony after MERGE MX
Homozygous for Hsa1, Hsa2 & Hsa3
Next select with pCas9-sgRNA Sc-α6,α7
Homozygoous for all 5 humanized loci
Heterozygous diploid
scl1Δ (α1)::HsPSMA6 (2N)
pre8Δ (α2)::HsPSMA2 (2N)
pre9Δ (α3)::HsPSMA4 (2N)
PUP5 (α6) / pre5Δ::HsPSMA1
PRE10 (α7) / pre10Δ::HsPSMA3
Sporulate
Homozygous diploid
SD -LEU + Thialysine
Mat α selection
Sporulation defect
G
Humanized α123
Mat A
with pCas9-sgRNA Sc-α1,α2,α3
Mate
Humanized α45
Mat α
Humanized α123 (homozygous)
Humanized α1234 (homozygous)
Humanized α12345 (homozygous)
Diploid
pre6Δ::HsPSMA7 (2N)
pre5Δ::HsPSMA1 (2N)
pre10Δ::HsPSMA3 (2N)
Genotype randomly picked colony after MERGE°
H
pCas9-sgRNA Sc-α7
Humanized α123467
Day 3
Day 5
Day 8
4 of 4 colonies show homozygous humanized loci for Hsa7 locus
Viable genotype
Inviable genotype
X

Figure S12. MERGE efficiently tests several higher-order combinations (>3) of humanized yeast α-proteasome core.

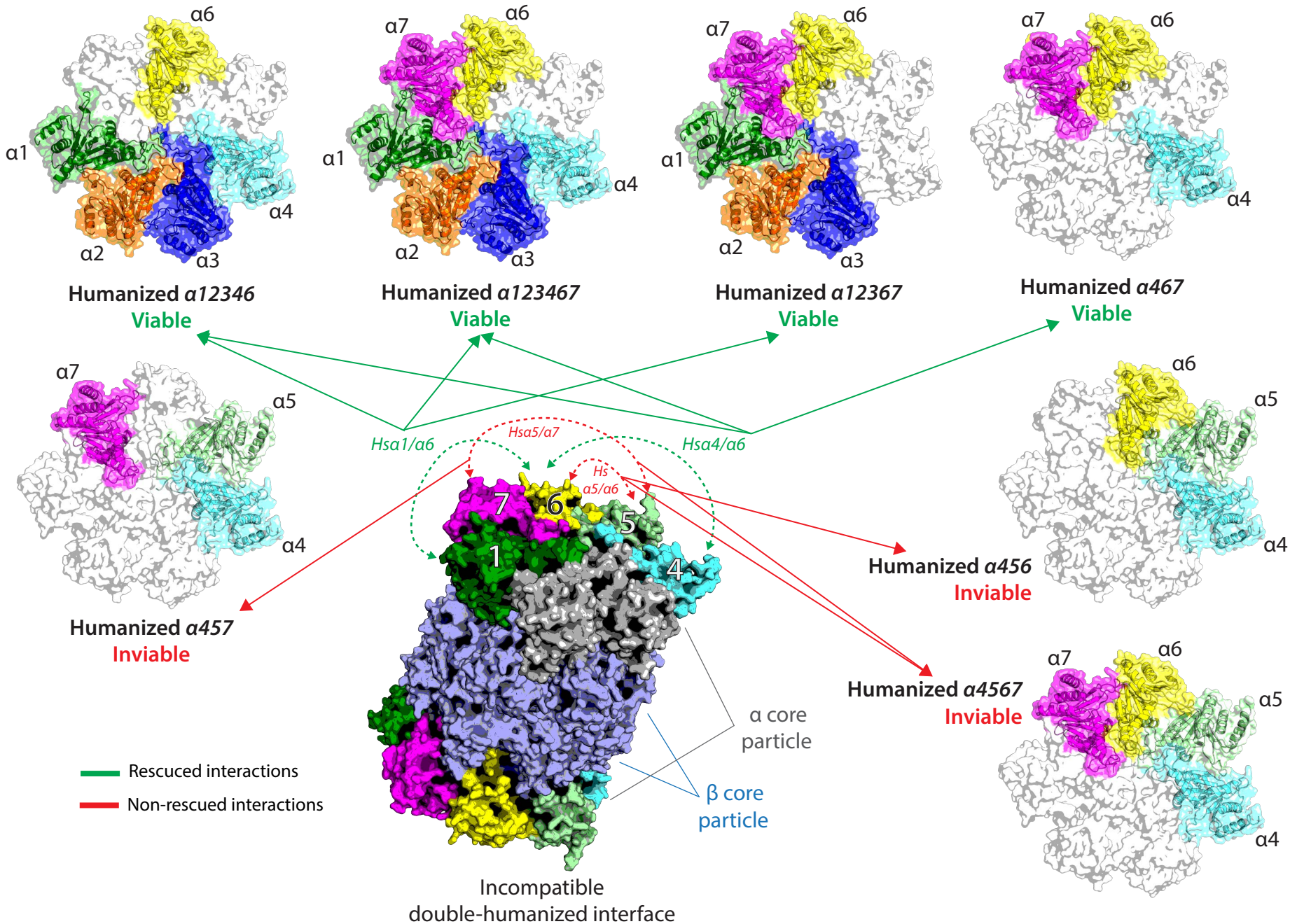
(**A**) Double-humanized haploid *Hs-α4α5* (Mat α) and *Hs-α6α7* (Mat α) strains were mated. The resulting heterozygous diploid strain was transformed with quadruple-sgRNA CRISPR reagent (**pCas9-sgRNA**$^{Sc\alpha4,\alpha5,\alpha6,\alpha7}$) to perform **MERGE**^{MX}. The transformation caused lethality with few survivors (**CFU_O/CFU_E =~0**). The survivors showed the conversion of only 2 of 4 loci (*Hs-α4* and *Hs-α5*), suggesting an incompatible quadruple-humanized genotype. However, the transformation of triple-sgRNA CRISPR reagent (**pCas9-sgRNA**$^{Sc\alpha4,\alpha6,\alpha7}$) in this strain was viable (**CFU_O/CFU_E =~1**). The genotyping of randomly picked colonies showed the conversion of 3 yeast to human loci (*Hs-α4, Hs-α6,* and *Hs-α7*), suggesting a triple-humanized genotype is viable with a heterozygous yeast-human *α5* locus. **(B)** Mating quintuple-humanized haploid *Hs-α1α2α3α4α7* [Mat A, also harboring a triple-sgRNA CRISPR reagent (**pCas9-sgRNA**$^{Sc\alpha1,\alpha2,\alpha3}$)] and single-humanized *Hs-α6* (Mat α) strains simultaneously combined three yeast to humanized *Hsα1, Hsα2, Hsα3* loci in a diploid yeast (**CFU_O/CFU_E =~1**). The strain harbors heterozygous yeast-human alleles for *α4, α6,* and *α7* loci. Transformation of the diploid strain with **(C) pCas9-sgRNA**$^{Sc\alpha4}$ or **(D) pCas9-sgRNA**$^{Sc\alpha6}$ generated viable quadruple-humanized strains (Hs-*α1α2α3α4* and Hs-*α1α2α3α6*) (**CFU_O/CFU_E =~1**). Genotyping of randomly picked colonies confirmed the homozygosity of the loci. **(E)** Using the quadruple-humanized strain *Hs-α1α2α3α4* as a background, the transformation of **pCas9-sgRNA**$^{Sc\alpha6}$ yielded a viable quintuple-humanized (Hs-*α1α2α3α4α6*) genotype (**CFU_O/CFU_E =~1**) confirmed using locus-specific PCR. **(F)** Double-humanized haploid *Hs-α6α7* (Mat α) and triple-humanized *Hs-α1,α2,α3* [Mat A, also harboring a triple-sgRNA CRISPR reagent (**pCas9-sgRNA**$^{Sc\alpha1,\alpha2,\alpha3}$)] strains were mated. **MERGE**^{MX} simultaneously combined three yeast to humanized *Hsα1, Hsα2, and Hsα3* loci in the diploid background (**CFU_O/CFU_E =~1**). The next transformation, using **pCas9-sgRNA**$^{Sc\alpha6,\alpha7}$ also homozygosed the *Hsα6* and *Hsα7* loci, generating a quintuple-humanized strain (**CFU_O/CFU_E =~1**). However, post sporulation, the mix plated on Mat α selection failed to produce any viable humanized haploid strains. **(G)** Triple-humanized *Hs-α1α2α3* [haploid Mat A, also harboring a triple-sgRNA CRISPR reagent (**pCas9-sgRNA**$^{Sc\alpha1,\alpha2,\alpha3}$)] and double-humanized *Hs-α4α5* (haploid Mat α) strains were mated generating a heterozygous diploid strain for 5 loci. The resulting diploid strain simultaneously combined three yeast to humanized *Hsα1, Hsα2, and Hsα3* loci while harboring two yeast-human heterozygous loci (*α4* and *α5)* (**CFU_O/CFU_E =~1**). The subsequent **MERGE**^{0} strategy (using **pCas9-sgRNA**$^{Sc\alpha4}$) generated a viable quadruple-humanized strain (*Hs-α1α2α3α4*) with heterozygous yeast-human *α5* locus (**CFU_O/CFU_E =~1**). The locus-specific PCR confirmed the homozygosity for most of the colonies. However, the subsequent CRISPR selection (using **pCas9-sgRNA**$^{Sc\alpha5}$) did not produce a viable strain (**CFU_O/CFU_E =~0**), suggesting an incompatible quintuple-humanized (*Hs-α1α2α3α4α5*) genotype. **(H)** Using the quintuple-humanized diploid strain *Hs-α1α2α3α4α6* harboring a yeast-human heterozygous *α7* locus as a background, the transformation of **pCas9-sgRNA**$^{Sc\alpha7}$ yielded a viable sextuple-humanized (Hs-*α1α2α3α4α6α7*) genotype (**MERGE**^{0}**, CFU_O/CFU_E =~1**) after incubation for 8 days. The genotype of the viable colonies confirmed the homozygosity and humanization of the *α7* locus.

Figure S13


A
Hsα7
Hsα6
Scα5
Hsα1
Scα4
Hsα2
Hsα3
Mat α
Haploid
Mate
+
Scα7
Scα6
Hsα5
Scα1
Hsα4
Scα2
Scα3
Mat A
Haploid
Human-yeast α1234567
(heterozygous)
Diploid
SCL1 (a1) / scl1Δ::HsPSMA6
PRE8 (a2) / pre8Δ::HsPSMA2
PRE9 (a3) / pre9Δ::HsPSMA4
PRE6 (a4) / pre6Δ::HsPSMA7
PUP2 (a5) / pup2Δ::HsPSMA5
PRE5 (a6) / pre5Δ::HsPSMA1
PRE10 (a7) / pre10Δ::HsPSMA3
SCL1locus
PRE8 locus
PRE9 locus
PRE6 locus
PUP2 locus
PRE5 locus
PRE10 locus
Sc
Hs
Sca1
Hsa1
Sca2
Hsa2
Sca3
Hsa3
Sca4
Hsa4
Sca5
Hsa5
Sca6
Hsa6
Sca7
Hsa7
B
MERGE°
CFU_E
pCas9 alone
CFU_O
pCas9-sgRNA^Sc-a1,a2,a3
Viable genotype
CFU_O/CFU_E = ~ 1
Genotype randomly picked
colony after MERGE^MX
SCL1locus
PRE8 locus
PRE9 locus
Sc
Hs
Hsa1
Hsa2
Hsa3
CFU_O
pCas9-sgRNA^Sc-a4,a5,a6,a7
Inviable genotype
CFU_O/CFU_E = ~ 0
X

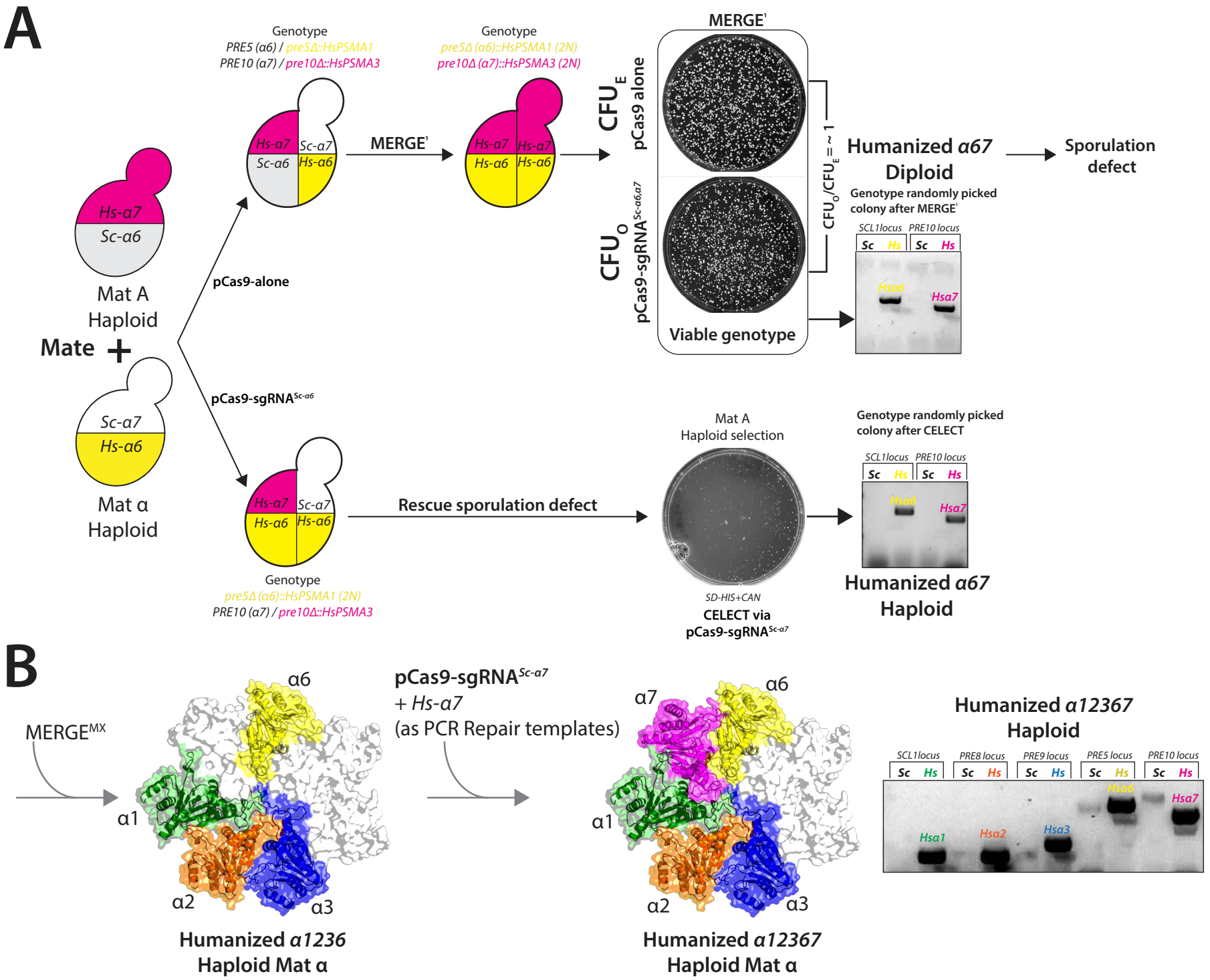
754 **Figure S13. MERGE enables the verification of the entire human α-proteasome core in**
755 **yeast.**
756 **(A)** Mating quintuple-humanized haploid *Hs-α1α2α3α6α7* (Mat α) and double-humanized *Hs-*
757 *α4α5* (Mat A) strains generated a diploid strain harboring yeast-human heterozygous alleles for 7
758 *α*-proteasome loci. Locus-specific PCR confirmed the heterozygous nature of the loci. **(B)** The
759 resulting diploid strain was transformed with a triple-sgRNA CRISPR reagent (**pCas9-**
760 **sgRNA**$^{Sc\alpha1,\alpha2,,\alpha3}$), resulting in a viable genotype homozygous for triple-humanized Hs-*α1α2α3* loci
761 (**CFU_O/CFU_E =~1**) with yeast-human heterozygous alleles for *α4, α5, α6,* and *α7* loci. The
762 subsequent transformation of a quadruple-sgRNA CRISPR reagent (**pCas9-sgRNA**$^{Sc\alpha4,\alpha5,\alpha6,\alpha7}$) in
763 the diploid strain was lethal, suggesting an incompatible genotype (**CFU_O/CFU_E =~0**).

Figure S14


α6
α1
α4
α2
α3
Humanized *α12346*
Viable
α7
α6
α1
α4
α2
α3
Humanized *α123467*
Viable
α7
α6
α1
α2
α3
Humanized *α12367*
Viable
α7
α6
α4
Humanized *α467*
Viable
α7
α5
α4
Humanized *α457*
Inviable
Hsα1/α6
Hsα5/α7
Hsα4/α6
Hs α5/α6
7
6
5
1
4
α6
α5
α4
Humanized *α456*
Inviable
α7
α6
α5
α4
Humanized *α4567*
Inviable
α core particle
β core particle
Rescued interactions
Non-rescued interactions
Incompatible double-humanized interface

764 **Figure S14. Inviable double-humanized genotypes are viable when neighboring**
765 **interactions are restored.**
766 Inviable double-humanized genotypes, *Hsα1/α6* and *Hsα4/α6,* become viable when the
767 neighboring interactions are restored in higher-order humanized strains (indicated as green lines).
768 However, incompatible paired genotypes comprising *Hsα5/α6* and *Hsα5/α7* could not be rescued
769 (marked as red lines). Proteasome core structures were generated using Pymol and PDB-1RYP.
770 Colored structures show humanized α yeast subunits.

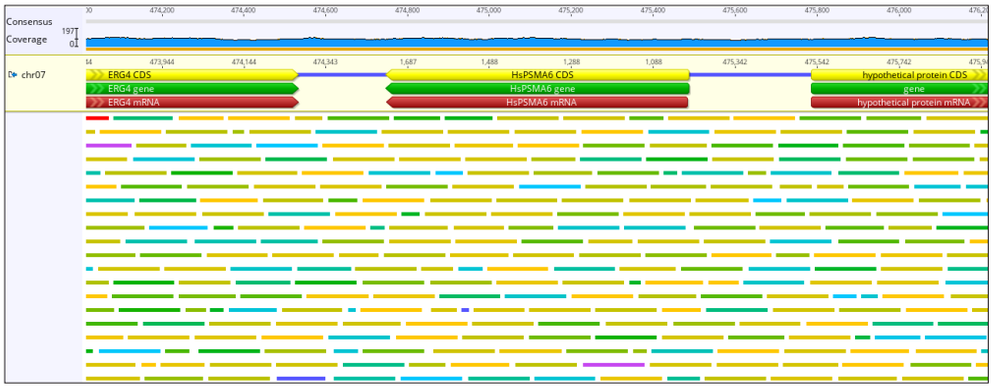
Figure S15


A
Genotype
PRE5 (α6) / pre5Δ::HsPSMA1
PRE10 (α7) / pre10Δ::HsPSMA3
Genotype
pre5Δ (α6)::HsPSMA1 (2N)
pre10Δ (α7)::HsPSMA3 (2N)
MERGE¹
CFU_E
pCas9 alone
CFU_O
pCas9-sgRNA^Sc-α6,α7
Viable genotype
CFU_O/CFU_E = ~ 1
Humanized α67
Diploid
Genotype randomly picked colony after MERGE¹
SCL1locus
PRE10 locus
Sc
Hs
Hsα6
Hsα7
Sporulation defect
Hs-α7
Sc-α6
Mat A
Haploid
Mate +
Sc-α7
Hs-α6
Mat α
Haploid
pCas9-alone
pCas9-sgRNA^Sc-α6
Rescue sporulation defect
Mat A
Haploid selection
SD-HIS+CAN
CELECT via
pCas9-sgRNA^Sc-α7
Genotype randomly picked colony after CELECT
Humanized α67
Haploid
B
MERGE^MX
α1
α2
α3
α6
Humanized α1236
Haploid Mat α
pCas9-sgRNA^Sc-α7
+ Hs-α7
(as PCR Repair templates)
α7
Humanized α12367
Haploid Mat α
Humanized α12367
Haploid
SCL1locus
PRE8 locus
PRE9 locus
PRE5 locus
PRE10 locus
Hsα1
Hsα2
Hsα3
Hsα6
Hsα7

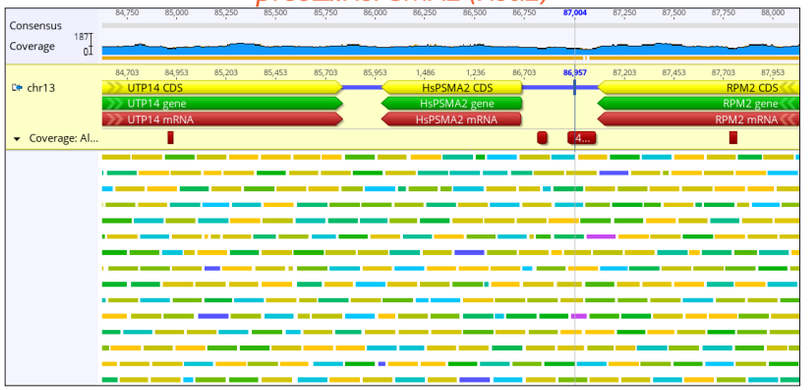
771 **Figure S15. Strategies to generate viable humanized haploid yeast strains when the**
772 **corresponding diploids manifest sporulation defects.**
773 **(A)** Mating single-humanized *Hs-α6* (Mat α, with **pCas9 alone** or **pCas-sgRNA**$^{Sc\text{-}\alpha6}$) and *Hs-α7*
774 (Mat A) strains generated two diploid strains; one harboring yeast-human heterozygous allele at
775 *α*6 locus and another with homozygous humanized *α6* locus. **MERGE**^{0} (using strategy **pCas-**
776 **sgRNA**$^{Sc\text{-}\alpha7}$) converted the heterozygous yeast-human *α7* locus to a homozygous human allele
777 at ~100% efficiency (**CFU_O/CFU_E =~1**). Locus-specific PCR confirmed the humanized
778 homozygous alleles. However, the diploid strains cannot sporulate. However, keeping the *α7*
779 locus as heterozygous rescues the sporulation defect. The haploid-specific selection (SD-
780 HIS+CAN; Mat A) allowed haploid yeast to grow, followed by **CELECT** (using **pCas-sgRNA**$^{Sc\text{-}\alpha7}$)
781 or simply genotyping the haploids to identify homozygous double-humanized Hs-*α6α7* genotype.
782 **(B)** The diploid quintuple-humanized Hs-*α1α2α3α6α7* strain shows a sporulation defect as in
783 **Figure S12F**. However, using a sequential CRISPR strategy, a haploid quadruple-humanized Hs-
784 *α1α2α3α6* strain enables the generation of haploid quintuple-humanized Hs-*α1α2α3α6α7* strain
785 to be used for subsequent **MERGE** experiments.

Figure S16

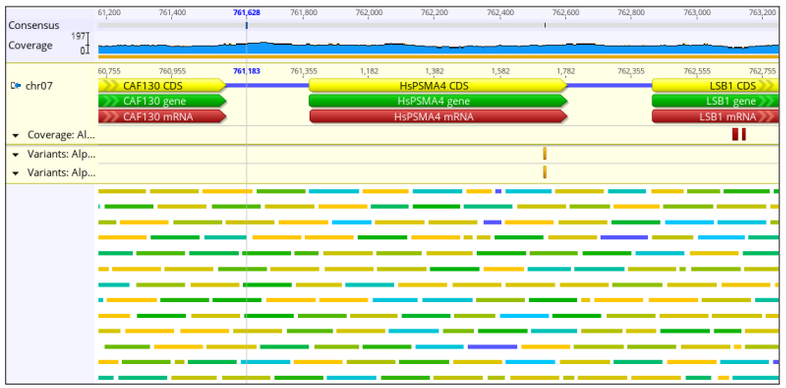
scl1Δ::HsPSMA6 (Hsα1)



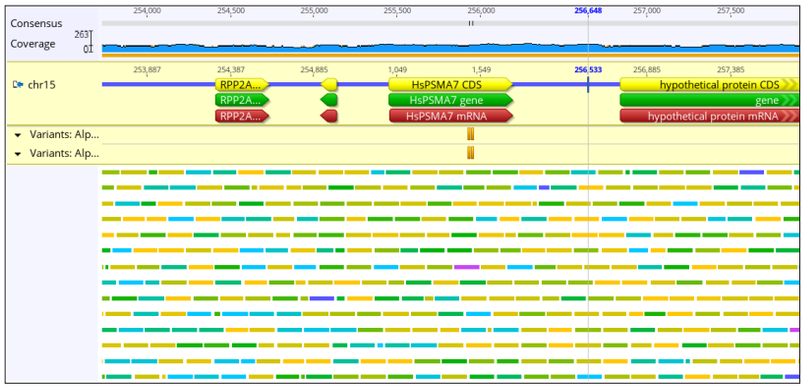
pre8Δ::HsPSMA2 (Hsα2)



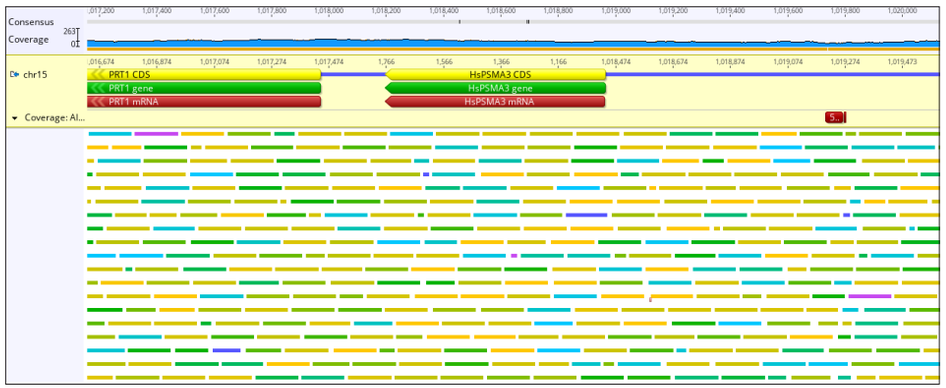
pre9Δ::HsPSMA4 (Hsα3)

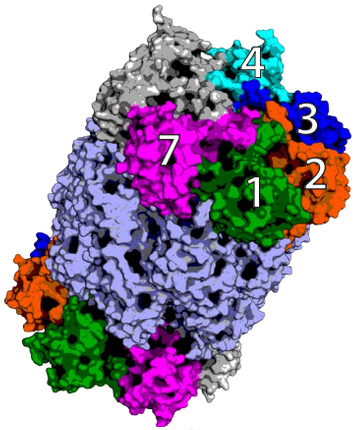


pre6Δ::HsPSMA7 (Hsα4)



pre10Δ::HsPSMA3 (Hsα7)





Humanized *α12347*
Quintuple-Humanized

786 **Figure S16. Whole-genome sequencing (WGS) of quintupled-humanized *Hs-α1α2α3α4α7***
787 **strain shows humanized loci alignment within the yeast genome.**
788 WGS analysis of quintupled-humanized *Hs-α1α2α3α4α7* shows humanized loci aligned to a
789 reference S288C genome with replaced human coding sequences. The human genes reside on
790 different yeast chromosomes, replacing their corresponding native yeast loci (Chr. 7, 13, and 15).
791 Green lines indicate the read lengths with a 200 base range, yellow is below, and blue is above
792 the range.

793 **Supplement for:**

794 Rapid, scalable, combinatorial genome engineering by Marker-less 795 Enrichment and Recombination of Genetically Engineered loci 796 (MERGE)

797
798
799

800 **Authors:**

801 Mudabir Abdullah[1], Brittany M. Greco[1], Jon M. Laurent[2†], Michelle Vandeloo[1], Edward M.
802 Marcotte[3], Aashiq H. Kachroo[1,‡]

803

804 **Affiliation:**

805 [1]Centre for Applied Synthetic Biology, Department of Biology,7141 Sherbrooke St. W,
806 Concordia University, Montreal, QC, Canada

807

808 [2]Institute of Systems Genetics, NYU Langone Health, New York, USA

809

810 [3]Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX, USA

811
812
813

814 [†]Current address: Pandemic Response Lab, 45-18 Ct Square West, Long Island City, NY 11101

815 [‡]Correspondence to - *aashiq.kachroo@concordia.ca*

816
817
818
819
820
821
822
823
824

825 **Materials & methods**
826
827 ***1. Targeting sgRNA design***
828 CRISPR-Cas9 targeting sequences consist of a 20 bp recognition sequence preceding an 'NGG'
829 sequence motif (PAM). To design our targeting sequences, we used the built-in guide RNA design
830 tool in version 11 of the Geneious software[36], using a recent version of the yeast genome
831 (available from http://www.yeastgenome.org/strain/S288C/overview) to screen for possible off-
832 target sequences[13,14]. We chose two high-scoring guide sequences to target each gene, requiring
833 that they be near the 5' end of the gene so any NHEJ repair would likely result in an early
834 frameshift. The guides were ordered as complementary oligos with overhangs according to the
835 sgRNA template described in the Yeast Tool Kit[15] (**Table S1**).
836
837 ***2. Plasmid construction using Yeast Tool Kit (YTK)***
838 Entry vectors for each guide RNA sequence were constructed by the BsmbI-mediated Golden
839 Gate (GG) assembly (Thermo) into plasmid pYTK050 from the Yeast ToolKit[15]. The targeting
840 sequences were ordered (IDT) as two single-stranded DNA oligos with a complementary region
841 and unique overhangs according to the 'sgRNA' template described in Lee, ME *et al*[15]. Briefly, the
842 two oligos for each guide were annealed by slow cooling from 95°C to 4°C (1-5 hrs) in a thermal
843 cycler, and 1µl of the annealed product was added to the entry vector GG reaction. Entry Vectors
844 were sequence-verified using custom primers.'Transcriptional unit' (TU) vectors (pYTK095) were
845 constructed by a BsaI-mediated GG assembly (Thermo). The entry vector for a particular guide
846 RNA was combined with a left connector part plasmid (ConLX) and right connector part plasmid
847 (ConRX) into the AmpR-ColE1 (pYTK095) backbone plasmid[13,14]. To create single-sgRNA TU
848 vectors, we assembled the sgRNA entry vector with left connector 1 (ConL1, pYTK003) and right
849 connector E (ConRE, pYTK072). To create TU vectors for the multi-sgRNA knockout, an entry
850 vector was combined with the appropriate connectors to enable proper subsequence knockout
851 vector assembly (i.e. a triple-sgRNA vector would have one TU vector with ConL1 and R2, one
852 with ConL2 and R2, and the third with ConL3 and RE). 'Knockout' (KO) vectors were constructed
853 by BsmbI-mediated GG assembly (Thermo). The appropriate TU vectors were assembled into
854 the CEN6-URA3 backbone and the Cas9-TU1 vector to create a KO vector. Each GG reaction
855 was performed in a 10µl volume, with approximately 20 fmol of each starting DNA molecule, along
856 with 0.5µl each of the appropriate restriction enzyme, 0.5µl of T7 DNA ligase (NEB), and 1µl of
857 10X T4 ligase buffer (NEB). A thermocycler was used to cycle between 16°C and 37°C each for
858 (5 minutes). 3µl of the reaction volume was transformed into competent DH5α *E. coli* and plated
859 on appropriate selective media for the backbone (i.e. chloramphenicol for entry vectors, ampicillin
860 for TU vectors, and kanamycin for KO vectors). Each backbone plasmid is a 'GFP-dropout' vector,
861 so correct clones were selected by screening for non-fluorescent colonies when viewed by blue
862 light and verified by sequencing.
863
864 ***3. Cas9-TU1***
865 The Cas9-TU1 vector was constructed by BsaI GG assembly of YTK parts pYTK002 (ConLS),
866 011 (PGK1 promoter), 036 (Cas9 coding sequence), 055 (ENO2 terminator, 067 (ConR1), and
867 095 (AmpR-ColE1 GFP dropout backbone).
868

869 ***4. CEN6-URA3 expression vector***
870 The CEN6-URA GFP dropout backbone vector was constructed by BsaI GG assembly of YTK
871 parts pYTK008 (ConLS'), pYTK047 (GFP dropout), pYTK073 (ConLE'), pYTK074 (URA3),
872 pYTK081 (CEN6/ARS4), pYTK084 (KanR-ColE1 RFP dropout vector).
873
874 ***5. An alternate strategy to directly clone annealed sgRNAs in a yeast expression***
875 ***vector***
876 To perform faster direct sgRNA-Cas9 plasmid construction without requiring three independent
877 cloning strategies (as described in section ***2***), we constructed **pCas9-GFPdo** to directly clone the
878 annealed sgRNA primers into yeast expression vectors. The sgRNA expression unit from the
879 pYTK050 plasmid was amplified via PCR using primers with BsaI sites that generated overhangs
880 similar to ConL1 (forward primer) and ConRE (reverse primer) to clone in CEN6-URA or CEN6-
881 *KanMX* yeast expression vectors. Finally, using the GG reaction protocol, the sgRNA TU PCR
882 was assembled into the CEN6-*URA3* or CEN6-*KanMX* backbone with the Cas9-TU1 vector to
883 create a direct knockout (**pCas9-GFPdo**) vector.
884
885 ***6. An alternate strategy to directly clone multiple sgRNAs TUs in a yeast expression***
886 ***vector for multiplexing in MERGE^{MX}***
887 The YTK kit allows cloning multiple sgRNA expression cassettes in a yeast expression vector.
888 However, the strategy involves three cloning steps. To build >1 sgRNAs TUs directly into the
889 yeast expression vector without generating intermediate connector vectors, we designed the
890 **pCas9-GFPMX** (CEN6-*URA3*) vector. Primers were designed to PCR amplify the GFP
891 expression cassette (for *E. coli* expression, pYTK047) with GG enzyme sites to clone GFP (BsaI
892 for cloning the GFP insert, the sites are lost after GG cloning). The BsmBI sites, with overhangs
893 similar to ConLS and ConRE, designed in the primers, allow the PCR-based sgRNA TU cloning
894 (obtained by ligating annealed sgRNA primers in pYTK050) in CEN6-*URA3* yeast expression
895 vector (**Table S1**). The primers used to amplify the sgRNA expression units now harbor connector
896 overhangs to clone >1 sgRNA expression units in tandem, eliminating the need to individually
897 make each connector clone.
898
899 ***7. Repair design and construction***
900 Repair DNA was designed to be a linear DNA molecule that contained the human coding
901 sequence, from the start codon to stop, with at least 100bp of flanking homology to the yeast
902 genome immediately upstream and downstream of the native yeast start and stop codons (**Figure**
903 **S3B**). They were constructed by PCR using human ORFeome[37] or MGC[38] clones as templates,
904 using primers with long extensions providing the homology sequence. Repair template PCRs
905 were performed with Accuprime Pfx (Thermo) as multiple 100µl reactions according to the
906 manufacturer's protocol, combined and purified using the Zymo DNA Clean or Qiagen PCR
907 Cleanup and Concentrator -25 kit. Approximately 5µg of repair DNA was included in the
908 transformation mix during sequential replacement transformations.
909
910 ***8. Humanization procedure***
911 Yeast transformations were performed using the Frozen-EZ Yeast Transformation II Kit from
912 Zymo Research (Cedarlane). In short, cells were grown to mid-log phase, washed with kit Solution

913 1, resuspended in kit Solution 2 and either used directly to transform or frozen at -80 C for future
914 use. In general, we transformed approximately 1µg of the KO plasmid along with 5µg of repair
915 DNA, which yielded anywhere from single to dozens of colonies, depending on transformation
916 efficiency and the number of simultaneous molecules necessary. In general, nearly all screened
917 clones were successful. For initial single replacement in the wild-type strains, we tested two
918 sgRNA sequences for each gene. These were tested in parallel, together with a no-repair control
919 transformation to assay for sgRNA effectiveness. Successful sgRNAs show zero or very few
920 colonies in the no-repair control transformation (**CFU_O/CFU_E =~0, Figure 1A & S4A**).
921
922 ***9. Engineering Carotenoid genes at the landing pad loci***
923 Single-sgRNA CRISPR reagents targeting three landing pad loci (*511B*, *USERX*, *FGF20*)[39] were
924 generated as described in ***section 1***. The repair templates were generated as previously
925 described in ***section 7***. Specifically, carotenoid gene transcription units were amplified from
926 plasmids harboring the transcription units using the primers with 80bp homology to the landing
927 pad loci[40] (**Table S1**). Each landing pad locus was edited in the SGA strain (Mat A) to harbor a
928 carotenoid gene transcription unit (*511B* for *CrtE*, *USERX* for CrtI, and *FGF20* for CrtYB) driven
929 by strong constitutive yeast promoters. The transformation was carried out using the Zymogen
930 frozen EZ yeast transformation kit (Cedarlane), and colonies were confirmed by PCR using
931 primers specific to the landing pad or the carotenoid gene loci as described below (**Figure S4B**
932 & **Table S1**).
933
934 ***10. Clone verification***
935 Clones were initially screened by colony PCR using a rapid DNA isolation method and colony
936 PCR[41]. Forward primers for PCR screening were designed such that the upstream primer would
937 bind in the yeast genome approximately 150bp - 500bp upstream of the yeast ORF or the landing
938 pad loci, to ensure it was outside of the homology region used for repair. Two reverse primers
939 were designed for each locus, one binding into the yeast and the other within the human or
940 carotenoid gene sequence. To facilitate multiplex PCR screening, each pair amplifies different-
941 sized bands for the yeast and human or carotenoid genes. Following plasmid loss, clones were
942 further verified by directly sequencing a PCR product using Sanger sequencing.
943
944 ***11. Plasmid loss to perform sequential editing of loci***
945 Successful clones were subjected to a plasmid loss procedure to alleviate any stress incurred by
946 the constitutive expression of the Cas9 protein and allow further knockout plasmids to be
947 transformed. To avoid any potential stress-induced defects, we rarely used 5-FOA counter-
948 selection to force loss of the URA3 plasmid, instead relying on spontaneous plasmid loss. One
949 two methods were used to identify spontaneous plasmid loss. Clones were grown overnight in
950 YPD and then again either spread onto YPD (100µl of 1:1000 dilution) and replica plated on SC
951 -Ura and YPD or patch plated (typically 6-12 colonies) to both minimal media lacking uracil and
952 YPD. This procedure resulted in around 10-60% of colonies losing the plasmid (estimated). For
953 a plasmid harboring URA3 selection, the strains that lost the plasmid were identified using 5-
954 FOA selection.
955
956 ***12. Growth curves***

957 Cells were diluted to approximately 0.01-0.02 OD660 (~2-5 x10^5 cells) in 150µl YPD across 3-4
958 replicates. They were grown and read in a BioTek Synergy H1 or Sunrise Tecan microplate
959 reader, with continuous double orbital shaking at 30℃ or 37℃. Reads were taken while stationary
960 every 10-20 minutes, and experiments were run for at least 24-36 hours. For spot assays, strains
961 were grown overnight in YPD at 30℃ and spotted with serial dilutions on YPD agar. The plates
962 were incubated for 2-3 days at 23℃, 30℃ and 37℃.
963
964 ***13. CRISPR-based selection (CELECT) to identify and enrich unique genotypes***
965 To test the efficiency of CRISPR selection, a subset of engineered strains (Haploid Mat A) were
966 mixed as a pool. Each strain was cultured overnight in YDP individually and back-diluted into a
967 pool the next day at equal OD and grown to mid-log phase. The strains used in the mix were all
968 of the same mating types to avoid random mating events during the experiment. Competent cells
969 were prepared and transformed with various pCas9-sgRNA^{locus} plasmids. We performed the
970 experiments as biological and technical replicates using different genotypes in the mix. The same
971 competent cell mix was either transformed with pCas9 alone or with various **pCas9-sgRNA$^{WT\ locus}$**
972 vectors, each harboring a different yeast-specific sgRNA expression unit. Competent cells were
973 generated using Zymo Research Frozen EZ kit transformation protocol. Each transformation was
974 plated on SD-URA. pCas9 alone served as a positive control allowing the growth of all genotypes
975 in the mixture. **pCas9-sgRNA^{ADE2}** was used to visually observe and quantify the efficiency of the
976 selection. The plasmid will eliminate all the other genotypes except the *ade2Δ::kanMX* (red
977 colonies due to the disruption of the *Ade2* gene). pCas9-sgRNA^{RPT5} was used as a negative
978 control as the yeast *Rpt5* gene is present in all the strains such that yeast cells should survive
979 this transformation. Colonies from each Petri plate were picked randomly and genotyped by locus-
980 specific PCR to confirm the CRISPR reagent exclusively selected a unique genotype as described
981 in section ***10***. (**Figure 1C**).
982
983 ***14. Mating and MERGE^0 assays***
984 The mating assays were carried out in the following ways:
985 ***14.1.*** The engineered strains with single gene modification in BY4741 (Haploid Mat A)
986 backgrounds were mated with SGA-haploid strains harboring haploid specific markers (SGA Mat
987 A; *can1Δ::STE2pr_Sp_his5his3Δ1leu2Δ0ura3Δ0;* select on SD-His+Canavanine & SGA Mat α;
988 *lyp1Δ::STE3pr_LEU2 his3Δ1leu2Δ0ura3Δ0;* select on SD-Leu+Thialysine). Each parental strain
989 was transformed with two different empty vector plasmids with distinct selectable markers *URA3*
990 and *KanMX*, and mated diploids were selected on SD-Ura+G418. The haploid strains were grown
991 overnight in 5ml selection media (SC -Ura or YPD +G418). The strains were mixed in a rich
992 medium and incubated at 30°C for 3-4 hrs. with shaking. 500µl of the mated mixture was washed
993 with distilled water, and 10µl of the mix was plated at different dilutions on solid agar media using
994 SD -URA+G418 selection and incubated for 2- 3 days at 30°C. Plasmids were cured using a
995 previously described strategy in section ***11***. The heterozygous colonies were confirmed by PCR
996 genotyping using primers specific to the wild-type and engineered loci as described in section ***10***.
997 The confirmed heterozygous strains for all proteasome and carotenoid strains were transformed
998 with CRISPR plasmids targeting wild-type loci and converting the parent heterozygous strain into
999 homozygous diploid strain. Colonies were picked randomly from the plate and genotyped for

1000 homozygosity using the same set of primers, confirming the loss of yeast and the presence of
1001 humanized or carotenoid loci.
1002 ***14.2.*** Alternatively, instead of an empty vector, the engineered humanized or carotenoid strains
1003 were transformed with a CRISPR reagent for which the corresponding strain harbors a CRISPR-
1004 resistant locus. The subsequent mating with an opposite mating-type strain with an empty vector
1005 (pCas9 alone or a backbone plasmid) results in the loss of heterozygosity and conversion of single
1006 loci (**MERGE^0**).
1007 In each scenario, the number of colonies obtained on plates transformed with CRISPR reagent
1008 (**CFU$_O$**) were compared with empty vector transformations (**CFU$_E$**) to calculate the efficiency of
1009 **MERGE^0**. Each experiment was performed at least 3 times.
1010
1011***15. Sporulation of diploid and selection of haploid yeast strains***
1012 Diploid strains were sporulated in a media containing 0.1% potassium acetate and 0.005% zinc
1013 acetate for 4-7 days. For Tetrad dissection, spores were spun down at 5000rpm for 5 minutes,
1014 resuspended in 200µl of 20mg/ml Zymolyase and incubated at 37°C for 25-30 minutes. The mix
1015 was then incubated at -20°C to stop the reaction. Cells were thawed on ice, and 20ul of the mix
1016 was plated into YDP for Tetrad dissection. Tetrad Dissection was performed using Spore play
1017 (Singer Instruments for every engineered proteasome spore on YPD and then replicated on SGA
1018 selection. For carotenoid and humanized loci in SGA background strain, the sporulation mix was
1019 directly plated on SGA selection. Locus-specific PCR verified the haploid strains for the presence
1020 or absence of engineered loci.
1021
1022***16. Estimating the efficiency of MERGE^0 using yeast ADE2 locus as a readout***
1023 The yeast *ADE2* locus was used to estimate **MERGE^0** using a CRISPR-resistant allele
1024 *ade2Δ::kanMX* on the homologous chromosome as a repair template. The sgRNAs were designed
1025 using Geneious version 11 as described in ***section 1*** (**Table S1**)**.** The **pCas9-sgRNA^{ADE2}** vector
1026 showed ON-target activity causing lethality in wild-type haploid or diploid strains. Occasional
1027 colonies show a red colony phenotype suggesting mutations via NHEJ at the *ADE2* locus. To
1028 estimate **MERGE^0** efficiency, the **pCas9-sgRNA^{ADE2}** was transformed in a diploid hetKO
1029 *ADE2*/*ade2Δ::kanMX* obtained from the yeast "Magic Marker" hetKO collection[26]. To further verify
1030 that the red colonies observed after **MERGE^0** are not due to the mutations in the *ADE2* locus
1031 (NHEJ) and instead due to conversion to *ade2Δ::kanMX* locus (HDR), the heterozygous diploid
1032 (ADE2/*ade2Δ::kanMX*) strain transformed with a control (**pCas9 alone**) and **pCas9-sgRNA^{ADE2}**
1033 plasmid were sporulated followed by tetrad dissection (**Figure S5A**). Haploid spores were selected
1034 on YPD or YPD + G418 (200µg/ml). Spores harboring a control plasmid showed the expected 2 :
1035 2 (red : white phenotype) and 2 : 0 (G418 resistant, G418 sensitive) phenotype compared to the
1036 **pCas9-sgRNA^{ADE2}** transformed cells that show 4 : 0 (red, white) colonies and 4 : 0 (G418 resistant,
1037 G418 sensitive) colonies phenotype, respectively, suggesting the conversion to *ade2Δ::kanMX*
1038 rather than the mutation of *Ade2* locus. In contrast, the co-transformation of **pCas9-sgRNA^{ADE2}**
1039 and oligo as a repair template harboring (100bp homology, 5X molar excess than the plasmid) to
1040 the 5' and 3' UTRs of *ADE2* locus in haploid wild-type yeast cells, shows resulted in significantly
1041 fewer survivors (**Figure 2C**; % **CFU$_O$/CFU$_E$** = 21.6+/-SD; N=4). However, this method is still far
1042 less-efficient than **MERGE^0** (**Figure 2C**; ~100% vs 21.6%).
1043

1044***17. Using MERGE^0 to perform one-step gene essentiality assays in yeast***

1045 HetKO diploid strains for 7 *α* proteasome core genes were obtained from the yeast "Magic Marker"
1046 hetKO collection[26]. The strains were transformed with either a single-sgRNA CRISPR reagent
1047 targeting the corresponding yeast *α* proteasome genes or the empty vector control and selected
1048 on SD -URA+G418. CRISPR reagent transformed plates caused lethality in 6 of 7 *α* proteasome
1049 hetKO strains, except in the case of non-essential *α3*. A similar assay in the hetKO
1050 *ADE2*/*ade2Δ::kanMX* strain showed viable homozygous null cells for a non-essential *ADE2* locus.
1051

1052***18. Testing MERGE^0 efficiency for all yeast loci on Chromosome I***

1053 All the strains harboring heterozygous knockout diploid loci from the yeast "Magic Marker"
1054 collection on chromosome I were arrayed in a 96-well format[26] (**Figure S6A**). A2 is the locus close
1055 to the left telomere, F1 near the centromere, and G10 close to the right telomere (**Figures 2D &**
1056 **S6A**). The sgRNA targeting the *KanMX* cassette was designed using the strategy mentioned
1057 earlier (**Table S1**). Since the bacterial selection for the previously described (***section 4***) yeast
1058 shuttle vector harbors a *Kan^R* cassette identical to the *KanMX* cassette in the case of hetKO
1059 strains, we designed a new expression vector with *AmpR* selection. The plasmid showed a lethal
1060 phenotype in the strains harboring the *KanMX* allele (haploid or homozygous diploid) in the
1061 genome (**Figure 2D**). Approximately 1µg of **pCas-sgRNA^{KanMX}** was transformed for each
1062 transformation.
1063 First, we performed a pilot assay using 5 hetKO strains with genes located at various regions on
1064 chromosome I (**Figure S6B**). These strains showed no lethality (**CFU_o/CFU_E of ~1**). Each strain
1065 was inoculated in 800µl of YDP + His(50 mg/L) + G418 (200 µg/ml) overnight in 96-well format.
1066 The following day, cultures were back-diluted and grown to the mid-log phase. Competent cells
1067 were generated in a 96-well format using the Gietz yeast transformation protocol (Gietz, 2007).
1068 Each transformation mix was transformed with either 1µg of pCas9 alone or 1µg of the **pCas9-**
1069 **sgRNA^{KanMx}** plasmid. An equal amount of the transformed cells were plated onto SD-URA and
1070 SD-URA + G418. All of the yeast loci (except for A10- *CNE1*) on chromosome I show ***MERGE^0***
1071 mediated conversion of *KanMX* to wild-type yeast locus irrespective of the position of the gene
1072 along the chromosome (**CFU_O/CFU_E =~1**) and simultaneously lost the KanMX cassette (**Figure**
1073 **S6A**).
1074

1075***19. Performing MERGE^1 and MERGE^{MX} at 2 or more loci using humanized proteasome and***
1076 ***carotenoid strains***

1077 CRISPR plasmids expressing 2 or more sgRNA cassettes in tandem were designed and
1078 constructed as described in sections ***1 - 6***. Each single-humanized α proteasome (BY4741, Mat
1079 A) was mated with SGA strain (Mat α) as described in section ***14***. Heterozygous strains from each
1080 mating mix were confirmed by locus-specific PCR genotyping as in section ***10***. ***MERGE^0*** enabled
1081 the conversion to humanized loci. The next round of mating in SGA strain (Mat A) followed by
1082 **MERGE^0** generated humanized strains in both mating-type SGA backgrounds. Carotenoid strains
1083 were directly generated using SGA compatible strains described in **section *9,*** followed by
1084 **MERGE^0** to obtain strains in both mating-types.
1085 Each double-sgRNA CRISPR reagent was lethal in singly engineered strains (**Figure S7B**). Mating
1086 between each distinct single-engineered genotype combines the loci as heterozygotes.
1087 Transformation of 0.5-1µg of double-sgRNA CRISPR reagent in heterozygous diploid strains either

showed no lethality (viable genotype, **CFU_O/CFU_E =~1**) or a lethal phenotype (inviable genotype, **CFU_O/CFU_E =~0**) (**Figures 3B, 3C &S7C**). Homozygosity was confirmed by randomly picking colonies after **MERGE[1]** and performing a locus-specific PCR confirming the presence or absence of yeast or engineered genotype.

For ***MERGE^{MX}*** to simultaneously target >2 loci, the CRISPR reagents were generated as described in sections ***2-4*** and ***6***. The haploid strains harboring combined engineering loci were mated with either a wild-type or strains carrying a different set of engineered loci. The diploids were transformed with either an empty vector or a multiplex-sgRNA CRISPR reagent. The genotypes were confirmed by locus-specific PCR as described in section ***10***.

20. Performing MERGE$^{M\&M}$ to combine 2 loci from the mated mix of single-humanized proteasome genes.

Single-humanized α proteasome strains *Hsα1*, *Hsα3*, and *Hsα7* of both mating-types were allowed to mate randomly in a mix. Each strain was inoculated at 0.3 OD in 2 ml YPD overnight, then added to a mix at equal OD and incubated on a shaking incubator for 4- 6 hrs 30°C. 500µl of the mix was further cultured overnight at starting 0.3 OD, followed by competent cell preparation. The mixture was transformed with either an empty vector or double-sgRNA CRISPR reagent. Several colonies were observed post ***MERGE$^{M\&M}$*** for *Hsα1α7,* and Hs*α3α7* paired genotypes. PCR-based genotyping confirmed the combination as a homozygous double-humanized diploid.

21. Testing the potential of CRISPR reagent to target multiple loci using the Green Monster strain

To address the scalability of targeting multiple yeast loci, we used the Green Monster (GM) strain[17]. The sgRNA target sequence was designed in GFP (S65T) variant using Geneious version 11 as described in section ***1***. **pCas9-sgRNA^{GFP}** was constructed using the method of direct cloning strategy as described in section ***5*** using a backbone with *Amp^R* (*E. coli*) and *KanMX* (yeast) selection. The Mat α mating-type Green Monster was used *MATα lyp1Δ his3Δ1 leu2Δ0 ura3Δ0 met15Δ0 can1Δ::GMToolkit-α [CMVpr-rtTA NatMX4 STE3pr-LEU2]*. The OFF-target activity of **pCas9-sgRNA^{GFP}** was tested in a haploid WT strain with no *GFP* gene showing a **$CFUo/CFU_E$ =~ 1**. The ON-target activity was measured transformation in natively tagged C-terminal fusion of GFP to *BRO1* gene *(Bro1-GFP)* obtained from the yeast GFP collection[42].

To estimate the loss-of-function mutations at the GFP locus, cells from an empty vector and **pCas9-sgRNA^{GFP}** transformed GM yeast cells were subjected to flow cytometry (BD Accuri™ C6 Plus Flow Cytometer). Cells were grown in 96-well format in YDP overnight and back diluted in SC media with 10µg/ml of doxycycline for ~48 hrs. The culture was diluted 1 in 10 in water and passed through the flow cytometer. Using BY4741 yeast with no GFP as a control, the cells showed mostly background fluorescence as expected. GM cells transformed with an empty vector show high fluorescence as this strain harbors 16 GFP cassettes serving as a positive control. Finally, random colonies were picked and pooled from haploid GM cells transformed with **pCas-sgRNA^{GFP,}** and the mixture passed through the flow cytometer. The cells show little GFP expression similar to the wild type, suggesting ON-target NHEJ-mediated mutations at nearly all GFP loci (**Figures S8**).

***22. Performing MERGE^{MX}* to assemble an entire Carotenoid pathway in yeast using mating and sporulation cycles.**

1132 SGA strains of both mating-types harboring single-carotenoid transcription units at the landing pad
1133 loci were generated in sections ***9 and 14***. The strains were cultured independently overnight,
1134 followed by mixing at a similar 0.3 OD and cultured on a shaking incubator for 5-6 hours. A hole
1135 was made on the lid of the Eppendorf tube to allow aeration, and the mixture was incubated at
1136 30℃ for 12 -16 hrs. The formation of diploids was confirmed by light microscopy. Next, 500µl of
1137 culture was centrifuged at 3500 rpm for 5 minutes. The supernatant was removed and
1138 resuspended in 2 - 5ml of sporulation media, followed by incubation on a rotating shaker for 5-8
1139 days at room temperature. Next, 500 ul of sporulation mix was centrifuged at 3500 rpm for 5 min
1140 and treated with Zymolyase. After the first sexual cycle, the cells were centrifuged, washed with
1141 water, and resuspended in rich media YPD. The cycle was repeated to generate a second set of
1142 diploids. The mixture was incubated for 4-6 hours in a shaking incubator and incubated at 30℃ for
1143 2 days. The confirmation and appearance of most diploids were again confirmed with a light
1144 microscope, and the 100µl of culture was inoculated overnight to make competent cells. The
1145 competent cells were transformed with either an empty vector or a triple-sgRNA CRISPR reagent
1146 **pCas9-sgRNA**^{511B,USERX,FGF20}. For the most part, we followed a GFP monster protocol. However,
1147 we did not use haploid or diploid-specific selection[43]. Instead, CRISPR selection was sufficient to
1148 enrich unique combined genotypes.
1149
1150***23. Genomic DNA isolation & Whole Genome sequencing, and SNP analysis***
1151 Genomic DNA was purified from single-humanized *Hsα1* and quintuple-humanized
1152 *Hsα1,α2,α3,α4,α7* strains using the Monarch Nucleic acid purification kit according to the
1153 manufacturer's protocol (NEB). Spheroplasts were obtained before the genomic DNA extraction
1154 for a high-quality DNA prep. Wild type and engineered strains were sequenced using Illumina
1155 MiSeq 2x150 at 30x coverage using 150-bp paired-end reads. The Geneious Pro Software
1156 [PMID: 22543367] and its included tools were utilized for pairing paired-end sequences,
1157 trimming ends and adapters based on quality using the BBDuk tool, and reference mapping
1158 using the Geneious Read Mapper algorithm at medium-low sensitivity, iterated 5 times. Reads
1159 were mapped to a reference BY4741 strain (for the engineered strain, the same reference
1160 sequence was used, but replacing the sequences of the 5 engineered genes with their human
1161 ortholog to ensure alignment with the humanized loci reads) using the Geneious Prime
1162 Software 2020.2.4 algorithm. Mapping was run at medium sensitivity: word length to allow for
1163 matching was set to 18 with a maximum permitted mismatch of 20% of the read length,
1164 maximum mismatches per read were set to 20%, a minimum 80% overlap was required, and
1165 reads with errors were set to accurately be mapped to repeat regions - this was iterated 5
1166 times to give the final mapping. For analysis, low coverage regions of below 2 standard
1167 deviations from the mean were excluded from SNP-calling, and only SNPs with a variant
1168 frequency of 0.90 or higher were considered. SNPs were called with a minimum variant
1169 frequency of 0.25, and low coverage regions below 2 standard deviations from the mean were
1170 excluded. SNPs that were unique to the engineered strain (i.e. not in the mappings of the wild-
1171 type strain) were marked. To ensure SNPs were not introduced by off-target CRISPR effects,
1172 300bp up- and down-stream of each SNP was searched for CRISPR-gRNA alignment at 75%
1173 sequence similarity.
1174

1175 **References**

1176

1177 36. Kearse, M. *et al.* Geneious Basic: an integrated and extendable desktop software
1178 platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649
1179 (2012).
1180 37. Lamesch, P. *et al.* hORFeome v3.1: a resource of human open reading frames
1181 representing over 10,000 human genes. *Genomics* **89**, 307–315 (2007).
1182 38. MGC Project Team *et al.* The completion of the Mammalian Gene Collection (MGC).
1183 *Genome research* **19**, 2324–2333 (2009).
1184 39. Bourgeois, L., Pyne, M. E. & Martin, V. J. J. A Highly Characterized Synthetic Landing
1185 Pad System for Precise Multicopy Gene Integration in Yeast. *ACS Synth Biol* **7**, 2675–2685
1186 (2018).
1187 40. Mitchell, L. A. *et al.* Versatile genetic assembly system (VEGAS) to assemble pathways
1188 for expression in S. cerevisiae. *Nucleic Acids Res* **43**, 6620–6630 (2015).
1189 41. Lõoke, M., Kristjuhan, K. & Kristjuhan, A. Extraction of genomic DNA from yeasts for
1190 PCR-based applications. *Biotechniques* **50**, 325–328 (2011).
1191 42. Huh, W.-K. *et al.* Global analysis of protein localization in budding yeast. *Nature* **425**,
1192 686–691 (2003).
1193 43. Suzuki, Y. *et al.* The green monster process for the generation of yeast strains carrying
1194 multiple gene deletions. *J Vis Exp* e4072 (2012) doi:10.3791/4072.

1195