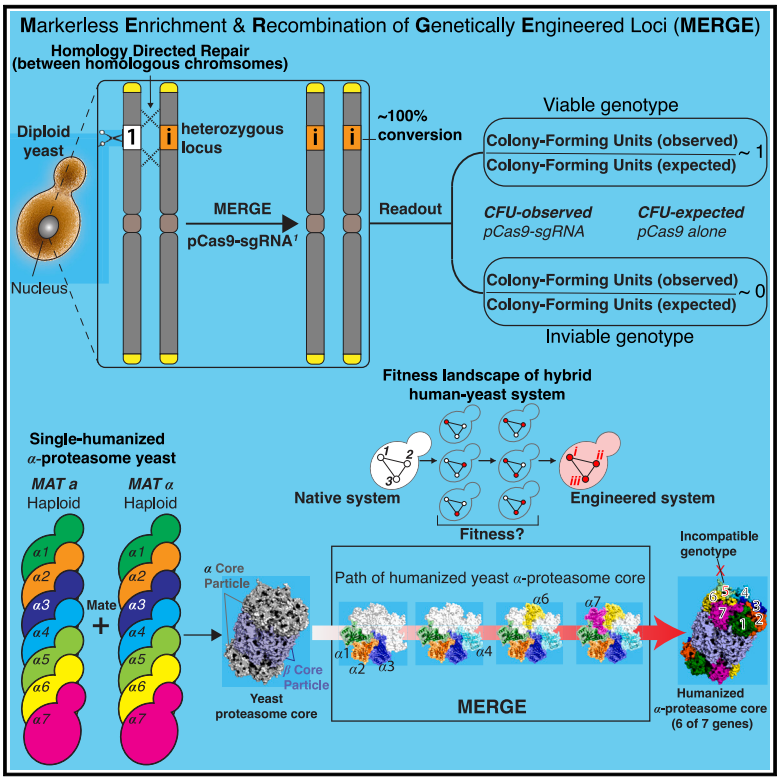
Cell Reports
Methods

Article

Rapid, scalable, combinatorial genome engineering by marker-less enrichment and recombination of genetically engineered loci in yeast

Graphical abstract



Highlights

- Cas9-induced gene drive efficiently converts heterozygous to homozygous yeast locus
- Yeast mating and Cas9 selection combines two or more separate edits into a single strain
- The method enables marker-less enrichment, recombination of genetically engineered loci
- MERGE reveals a fitness-driven path to humanize α-proteasome core subunits in yeast

Authors

Mudabir Abdullah, Brittany M. Greco, Jon M. Laurent, ..., Michelle Vandeloo, Edward M. Marcotte, Aashiq H. Kachroo

Correspondence

aashiq.kachroo@concordia.ca

In brief

Abdullah et al. use yeast mating and CRISPR-Cas9 selection to combine many individual genetic edits into a single strain. The method facilitates marker-less enrichment and recombination of genetically engineered loci (MERGE). MERGE engineers the entire genetic systems in yeast by exploring the fitness landscape.

Abdullah et al., 2023, Cell Reports Methods 3, 100464
May 22, 2023
https://doi.org/10.1016/j.crmeth.2023.100464

CellPress

Cell Reports Methods

CellPress
OPEN ACCESS

Article

Rapid, scalable, combinatorial genome engineering by marker-less enrichment and recombination of genetically engineered loci in yeast

Mudabir Abdullah,[1] Brittany M. Greco,[1] Jon M. Laurent,[2] Riddhiman K. Garge,[3] Daniel R. Boutz,[3] Michelle Vandeloo,[1] Edward M. Marcotte,[3] and Aashiq H. Kachroo[1,4,*]

[1]Centre for Applied Synthetic Biology, Department of Biology, Concordia University, 7141 Sherbrooke St. W, Montreal, QC, Canada
[2]Institute of Systems Genetics, NYU Langone Health, New York, NY, USA
[3]Department of Molecular Biosciences, The University of Texas at Austin, Austin, TX, USA
[4]Lead contact
*Correspondence: aashiq.kachroo@concordia.ca
https://doi.org/10.1016/j.crmeth.2023.100464

MOTIVATION Engineering entire biological processes in yeast, controlled for expression and genomic context, is challenging due to inefficient homology-directed repair (HDR) using linear exogenous repair templates, often resulting in non-quantifiable readouts. Thus, an inability to obtain the edit can be attributed to inefficient HDR or inviable genotype. Therefore, accomplishing multiplexed and scalable genomic editing necessitates compatible technology. We developed a CRISPR-Cas9-based combinatorial genome editing method (MERGE) that facilitates a fitness-driven combination of all individual genetic edits, eliminating the need for selection markers while using a quantitative readout of colony-forming units.

SUMMARY

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 a precise and multi-site genome editing approach that combines all edits without selection markers using CRISPR-Cas9. 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 marker-less enrichment and recombination of genetically engineered loci (MERGE). We show that MERGE converts single heterologous loci to homozygous loci at ~100% efficiency, independent of chromosomal location. Furthermore, MERGE is equally efficient at converting and combining multiple loci, thus identifying compatible genotypes. Finally, we establish MERGE proficiency by engineering a fungal carotenoid biosynthesis pathway and most of the human α-proteasome core into yeast. Therefore, MERGE lays the foundation for 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] In addition, 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* and yeast genomes using yeast HR.[6–9]

Despite 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*[10] (e.g., multiplex automated genome engineering), 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 in the case of efforts to systematically humanize yeast genes,[11,12] 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.[13–17] However, the approach gets progressively more challenging for multi-gene systems when the fitness of the intermediate genotypes is unknown. 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.[18] Using an SGA-like strategy to build a heterologous multigene 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.[19] However, implementing the GM strategy can be challenging for heterologous pathway engineering, with an added burden of expressing fluorescent cassettes.[20] Thus, accomplishing large-scale combinatorial genome editing in yeast necessitates genetic tools that circumvent these requirements.

This work describes a CRISPR-Cas9-based method to readily combine genetically engineered loci without requiring 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[21–24] 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

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

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.[25] Therefore, Cas9-single guide RNA (sgRNA)-induced lethality serves as a rapid test for a functional CRISPR plasmid (pCas9-sgRNA^{locus}) in yeast (colony-forming units observed sensitive [CFU_{OS}]) compared with the Cas9 alone (CFUs expected [CFU_E]; $CFU_{OS}/CFU_E = \sim0$). In contrast, a yeast strain that harbors a corresponding engineered locus resistant to further targeting allows the survival of colonies (CFUs observed resistant [CFU_{OR}]) similar to the pCas9 alone ($CFU_{OR}/CFU_E = \sim1$) (Figure 1A). Henceforth, for the sake of simplicity, we refer to CFU_{OS} and CFU_{OR} as CFUo.

We tested this strategy by humanizing α-proteasome core genes in yeast that are functionally replaceable by their human counterparts.[26] Using CRISPR-Cas9, we replaced each yeast α-proteasome gene with its human ortholog at the native loci (Figure S1). Additionally, we modified non-essential loci to test whether the CRISPR-mediated selection is broadly applicable and found that all engineered strains are resistant to the corresponding CRISPR plasmid-induced DSB ($CFU_O/CFU_E = \sim1$; Figures 1B, S2A, and S2B).

To further verify if the resistance of humanized strains is exclusive to its corresponding CRISPR plasmid, all single-humanized α-proteasome strains were mixed in culture (Figure 2A). The CRISPR plasmid that expresses only Cas9 serves as a "no selection control" (Figure 2B). To quantify the enrichment of unique genotypes, we inoculated an *ade2Δ::kanMX* haploid strain in the mixture. pCas9-sgRNA^{ADE2} exclusively enriched for resistant *ade2Δ* genotype (red colonies), whereas all other genotypes in the mix harboring a wild-type *ADE2* locus are inviable (Figure 2C). Conversely, the transformation of pCas9-sgRNA^{RPT5} targeting *RPT5* (a base subunit of the proteasome complex), for which all strains in the mix harbor a wild-type copy, shows no survivors (Figure 2D). Furthermore, we demonstrate that each CRISPR plasmid targeting yeast α-proteasome genes selected a corresponding humanized genotype from a mix, respectively (Figure 2E). 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 plasmid targeted to either allele should enable the conversion to homozygous diploid at high efficiency while enriching the desired genotype (MERGE level 0 [MERGE^0]) (Figure 3A). We used the *ADE2/ade2Δ::kanMX* heterozygous knockout (hetKO) diploid strain to quantify the efficiency of MERGE^0. The wild-type *ADE2* allele is susceptible to pCas9-sgRNA^{ADE2}-mediated DSBs. In contrast, the *ade2-Δ::kanMX* allele is resistant, providing a readout of conversion as the loss of function of *ADE2* results in a red color colony

CellPress
OPEN ACCESS

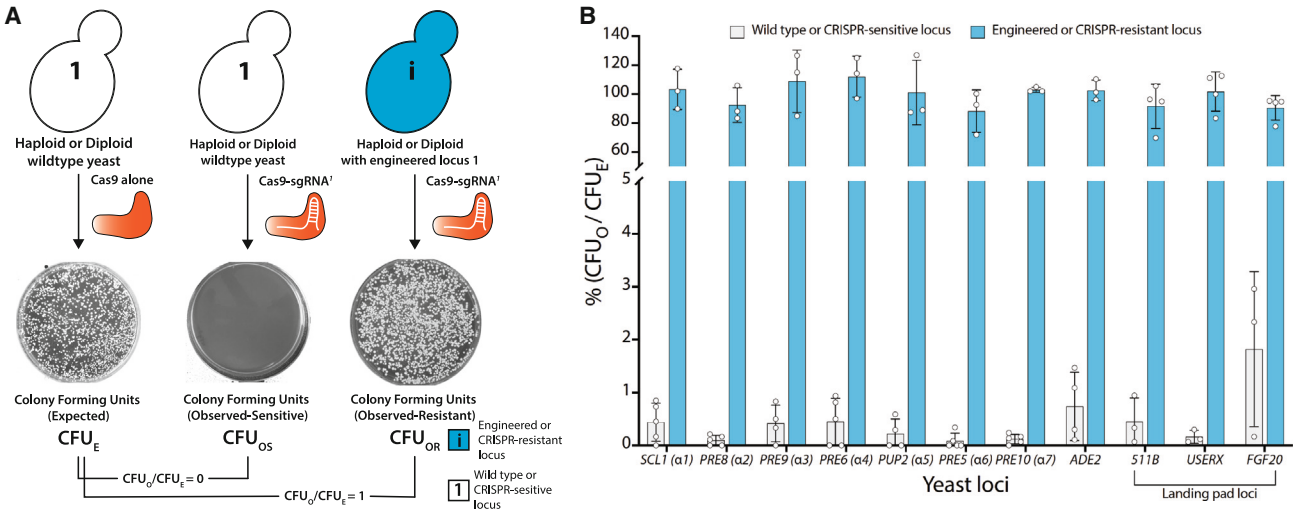


Figure 1. CRISPR-Cas9-mediated selection to enrich unique genotypes

(A) Schematic shows pCas9-sgRNA^{1} targeted to any yeast locus (locus 1) leads to lethality (colony-forming units observed sensitive [CFU_{OS}]) compared with the vector without sgRNA (pCas9 alone; [CFU_E]). However, the modification of locus 1 to i prevents targeting by the corresponding CRISPR plasmid (CFUs observed resistant [CFU_{OR}]). A relative CFU readout (CFU_O/CFU_E = ~1) identifies the modified genotypes.

(B) Each CRISPR plasmid targeted to several wild-type yeast loci shows near 100% lethality (gray bars, $\%CFU_O/CFU_E$ = ~100 or CRISPR-sensitive loci). However, after editing the loci (by singly humanizing yeast α-proteasome genes and replacing *ADE2* with a *kanMX* cassette [*ade2Δ::kanMX*] and inserting carotenoid genes at landing pad loci), the strains show resistance to the corresponding CRISPR plasmid (blue bars or CRISPR-resistant loci, $\%CFU_O/CFU_E$ = ~100). Data are shown as mean with standard deviation performed in triplicate.

phenotype. We show that the transformation of pCas9-sgRNA^{ADE2} in the *ADE2/ade2Δ*:kanMX hetKO strain shows resistance while also converting the locus to *ade2Δ*:*kanMX* at ~100% efficiency (Figures 3B and S3A). MERGE^{0} performs with a comparable proficiency at single-humanized α-proteasome and landing pad loci (Figures 3B, S2, S3B, and S3C). 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. A CRISPR plasmid targeting *kanMX* cassette (pCAS9-sgRNA^{kanMX}) (Figure S3D) in hetKO diploid strains converted all heterozygous loci to homozygous wild-type alleles, respectively, with simultaneous loss of *kanMX* cassette (except for *CNE1*) (Figures 3C, S4A, and S4B).

Similarly, Cas9-sgRNA targeting a wild-type yeast locus instead of the *kanMX* cassette in a hetKO diploid strain enables single-step gene essentiality screening in yeast, reducing the steps to test essentiality compared with the classical methods such as Tetrad dissection or SGA analysis.[27,28] Except for a non-essential *α3* strain, all six essential yeast α-proteasome strains are inviable post-transformation of the corresponding CRISPR plasmids (Figure S3E).

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 is efficient at converting two distinct loci (MERGE^{1}). Mating yeast strains, each with one modified locus, facilitate the combination as heterozygotes at two separate loci. Plasmids harboring Cas9 with two sgRNAs that target each unique locus should allow simultaneous conversion to paired CRISPR-resistant loci, respectively (Figure 4A). 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. To verify if MERGE^{1} can simultaneously convert two distinct loci, we tested the efficiency of combining paired-humanized $\alpha3/\alpha7$ and $\alpha1/\alpha7$ genotypes. We show that MERGE^{1} is highly efficient, as each heterozygous genotype is resistant to the double-sgRNA-CRISPR plasmid while converting and combining two humanized loci (Figures 4B, S5A). In comparison, the double-sgRNA-CRISPR plasmid transformation is lethal in wild-type and single-humanized strains (Figures 4B, S5A, and S5B). Additionally, although viable, the double-humanized *α1/α7* genotype shows a sporulation defect (a phenotype associated with all *Hsα7* strains) (Figure S5A). In contrast, a double-humanized *α5/α7* is inviable as a combined genotype, as evidenced by double-sgRNA-CRISPR plasmid-mediated lethality (Figure S5A). Thus, MERGE^{1} enriches and selects only compatible homozygous double-humanized strains without diploid-specific selection. Furthermore, while a single-humanized *Hsα6* strain is temperature sensitive (TS) at 37°C, a combination with the neighboring humanized *α7* gene rescues the phenotype (Figure S5C). We show that the TS phenotype is associated with a specific variant of *Hsα6* (variant 1, 37> glycine), whereas another common variant (variant 2, 37> valine) shows no growth defect at 37°C (Figure S5C). Thus, MERGE^{1} revealed the fitness of paired-humanized genotypes similar to synthetic genetic interactions without requiring genetically linked markers or diploid-/haploid-specific selections.[18]

To systematically determine if there are specific pairwise restrictions to the humanization of yeast α-proteasome subunits, we

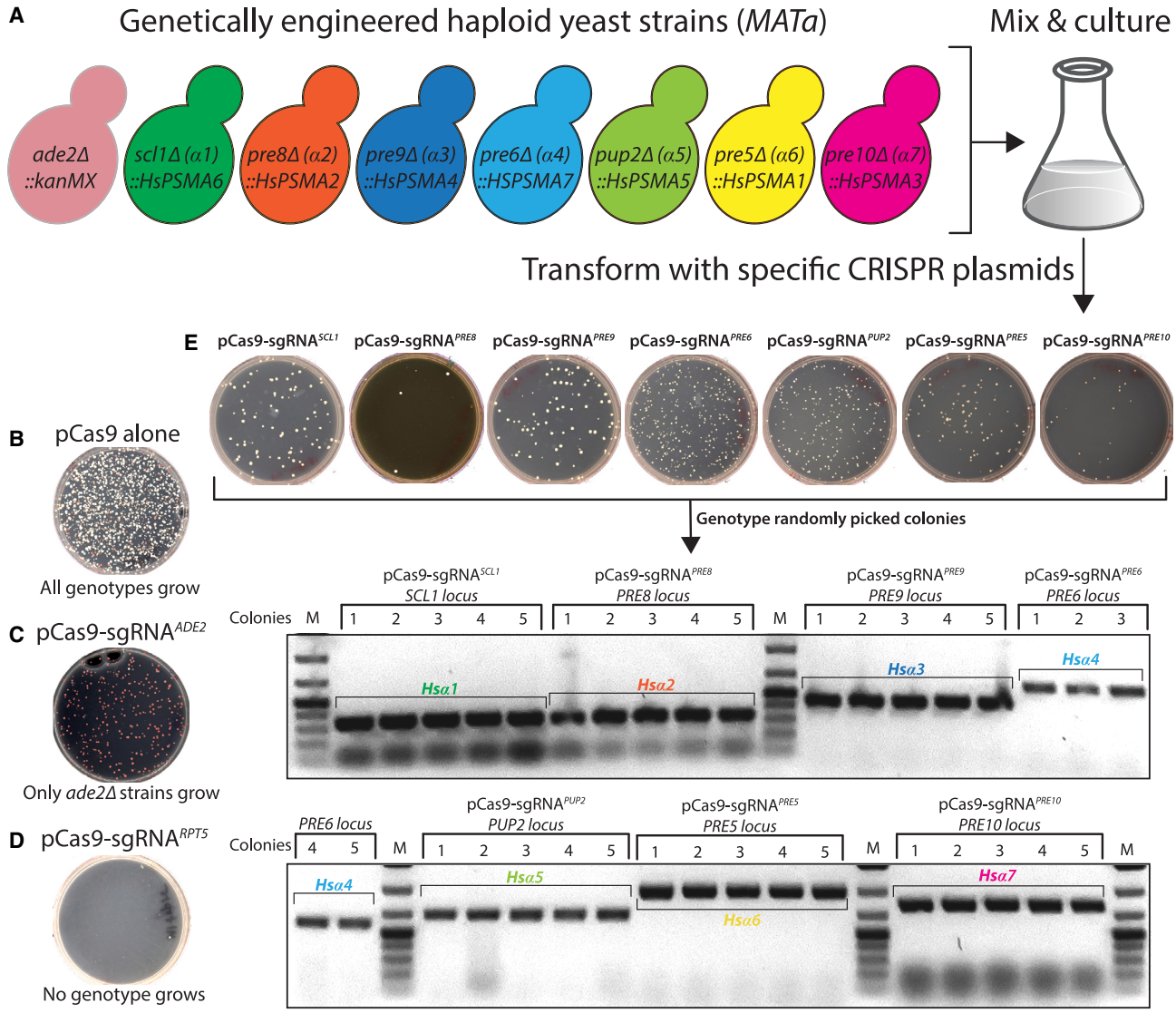


Figure 2. CRISPR-Cas9 selection mimics classical selection markers to enrich unique genotypes from a mixed culture

(A) Each unique genetically modified strain can be enriched from the mixture using the corresponding CRISPR reagent.

(B) The transformation of pCas9 alone serves as a control.

(C) Haploid *ade2Δ::kanMX* strain quantifies the efficiency of the CRISPR selection (pCas9-sgRNA^{ADE2}), enriching only the *ade2Δ* homozygous genotypes (red-colony phenotype).

(D) The transformation of pCas9-sgRNA^{RPT5} causes lethality as all the strains in the mix harbor a wild-type copy of the *RPT5* gene.

(E) Similarly, each CRISPR reagent specific to individual yeast α-proteasome genes also selected the corresponding humanized strains, as demonstrated by PCR-based genotyping of randomly picked colonies.

obtained all possible diploid double-humanized heterozygous genotypes (21 different genotypes). The corresponding CRISPR plasmid selections and a simple readout of CFU_O/CFU_E identified the permitted double-humanized genotypes (17/21 genotypes). The data reveal that only specific double-humanized genotypes are viable, whereas some are not (Figures 4C and S5D). The incompatibility of paired genotypes comprising *Hsα1,α6*; *Hsα4,α6*; *Hsα5,α6*; and *Hsα5,α7* may likely be due to the missing neighboring interactions within the α-proteasome core (Figure S5D′). In parallel, a sequential editing strategy successfully engineered some paired-humanized genotypes to confirm that MERGE[1] represents viable/fit paired genotypes (Figure S5E). However, it did not provide a clear perspective of incompatibilities due to the inability to successfully generate all genotypes.

Next, we explored the strategy to enrich double-humanized genotypes randomly from a mixture (mix, mate, and MERGE [MERGE$^{M\&M}$]) (Figure 4D). Culturing single-humanized *α1, α3*, and *α7* haploid strains of both mating types produced various heterozygous double-humanized genotypes in the mix. Each double-sgRNA-CRISPR plasmid selection enriched

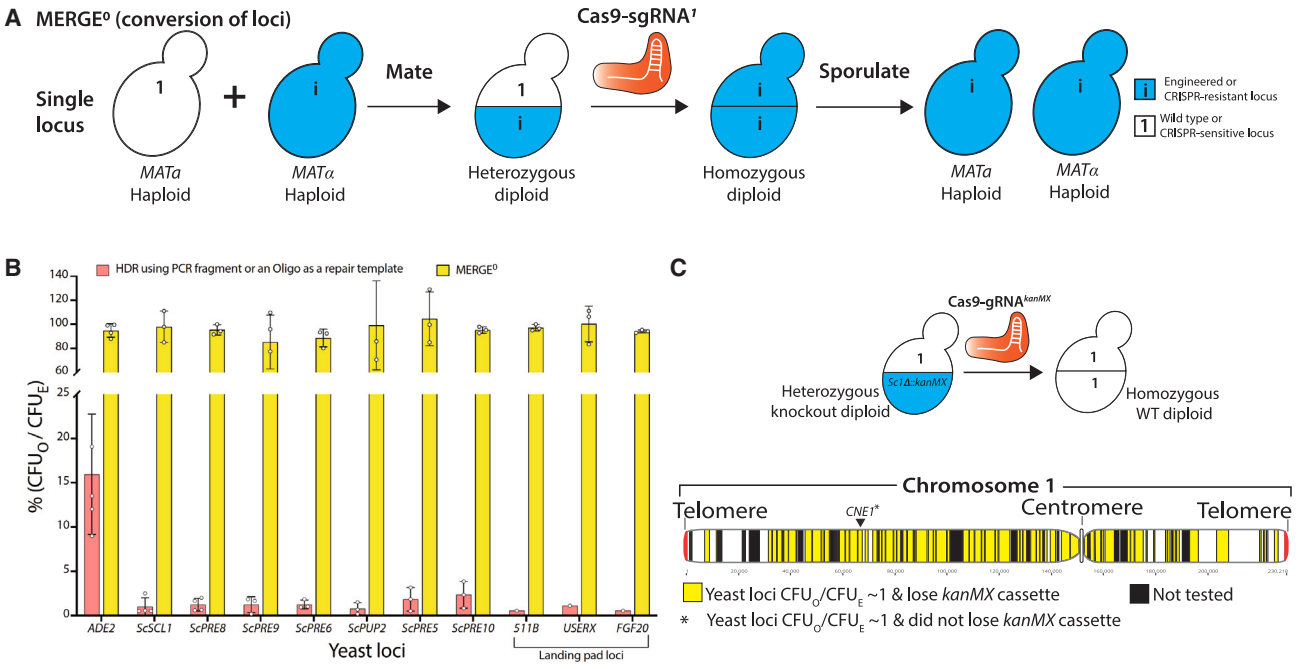


Figure 3. MERGE level 0 (MERGE^0) efficiently converts a single heterozygous to a homozygous locus

(A) Schematics showing mating haploid yeast strains, each harboring a different allele at a single locus, generates a heterozygous locus. CRISPR plasmid targeted to one of the alleles (locus 1, CRISPR-sensitive locus) initiates recombination using the homologous chromosome with a CRISPR-resistant locus (locus i) as a repair template. Sporulation of the resulting homozygous diploid yields the desired genotype in both mating types.

(B) CRISPR plasmids targeted to several wild-type yeast loci show near 100% lethality with few surviving colonies while simultaneously providing exogenous repair templates (oligo for *ADE2* locus and PCR fragments for the remaining loci) (red bars). However, each heterozygous diploid displays resistance to CRISPR plasmids, respectively (%CFU_O/CFU_E = ~100, yellow bars). Data are shown as mean with standard deviation performed in triplicate.

(C) Using heterozygous diploid knockout (hetKO) strains arrayed across the entire yeast chromosome I, MERGE^0 similarly converted every knockout allele to the wild-type locus (yellow = CFU_O/CFU_E = ~1), except in the case of *CNE1*.

the corresponding paired genotype from the mix while converting the wild-type yeast to humanized loci (Figures 4Dii and 4Diii). The transformation of pCas9-sgRNA$^{Sc\text{-}\alpha1,\alpha6}$, a selection for a non-existing genotype, does not yield any viable genotype (Figure 4Div). Thus, MERGE can be scaled to obtain several paired genotype combinations of engineered loci from a mix.

MERGE is scalable to combine multiple genetically engineered loci

The CRISPR-Cas9 system enables multiplexed editing by introducing multiple sgRNAs.[29] To test scalability, we designed MERGE^{MX} (MERGE, mate, and multiplex) to verify if >2 genetic loci can simultaneously convert to engineered loci by building the 4-gene carotenoid biosynthesis pathway from the carotenogenic yeast *Xanthophyllomyces dendrorhous* into Baker's yeast. The carotenoid pathway provides a colony color readout as a proxy for pathway engineering (Figure 5A). MERGE^0 generated haploid strains of opposite mating types for each carotenoid transcription unit, MERGE^1 provided the double-carotenoid genotypes (Figure 5B), and MERGE^{MX} enriched a complete homozygous carotenoid pathway genotype (Figure 5C). Furthermore, genotyping of randomly picked dark orange colonies confirmed the conversion and combination of the engineered loci (Figure 5C).

To test whether MERGE^{MX} can perform the combination of >2 humanized α-proteasome genotypes, we mated a previously obtained triple-humanized *Hs*α*1*α*2*α*3* strain with a wild-type strain generating a heterozygous diploid for three distinct loci. The triple-sgRNA-CRISPR plasmid (pCas9-sgRNA$^{Sc\text{-}\alpha1,\alpha2,\alpha3}$) converted all yeast loci to human versions (Figure 5D). By performing mating of various humanized strains followed by CRISPR plasmid selections, we explored several triple-humanized genotypes, successfully identifying compatible (Figures 5D and S6A) and inviable combinations (Figures S6B and S6C). Thus, if a triple-humanized intermediate genotype is viable, MERGE^{MX} can enrich and combine the specific genotype.

To further address throughput, the transformation of CRISPR plasmid targeting a *GFP* expression cassette in a GM strain[19] (16 *GFP* loci) resulted in few surviving colonies that failed to show GFP expression, likely due to mutations at *GFP* loci as a result of error-prone NHEJ repair, suggesting successful targeting of the majority of *GFP* cassettes (Figure S6D).

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

To explore the proficiency of MERGE for testing >3 combinations of engineered loci, we asked if an entire yeast heptameric α-proteasome core is humanizable. As an alternate strategy, we also tested the feasibility of sequential engineering using repetitive

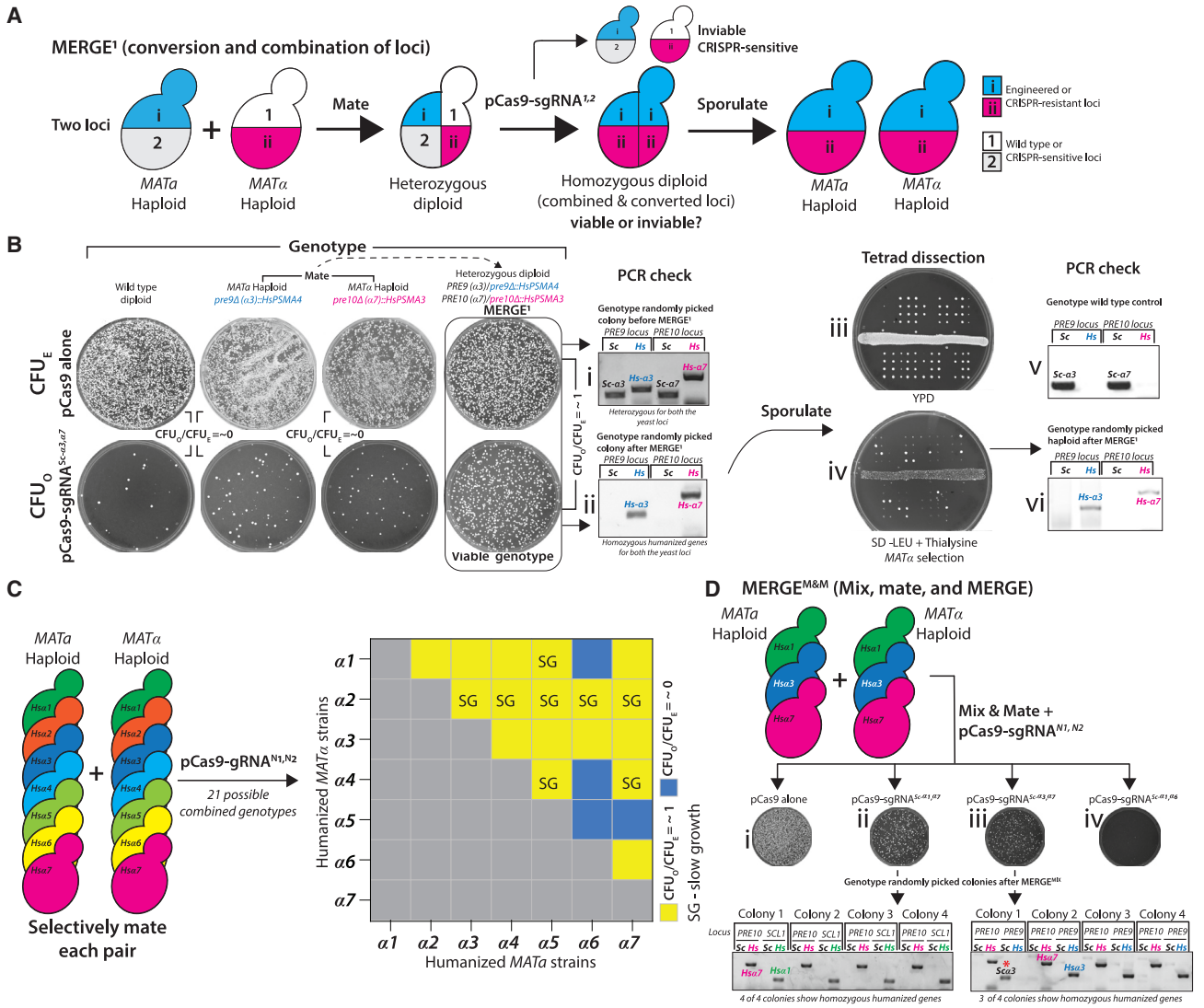


Figure 4. MERGE[1] efficiently converts and combines two heterozygous to homozygous loci

(A) The schematic shows the mating of haploid yeast strains, each harboring two different alleles. Double-sgRNA-CRISPR plasmid targeted to both the alleles (CRISPR-sensitive loci, 1 and 2) initiates recombination at both loci using the homologous chromosome with CRISPR-resistant loci (i and ii) as a repair template, thus enabling the simultaneous conversion of 2 loci. Sporulation of the resulting homozygous diploid yields the desired combined genotype in both mating types.

(B) Double-sgRNA-CRISPR plasmid (pCas9-sgRNA$^{Sc\text{-}\alpha3,\alpha7}$) targeted to two wild-type yeast loci (proteasome $\alpha3$ and $\alpha7$ genes) quantifies the efficiency of MERGE[1]. The transformation of pCas9-sgRNA$^{Sc\text{-}\alpha3,\alpha7}$ in the wild-type and single-humanized haploid *Hsα3* or *Hsα7* strains is lethal, with few surviving colonies ($CFU_O/CFU_E = \sim0$). However, the transformation of pCas9-sgRNA$^{Sc\text{-}\alpha3,\alpha7}$ in a heterozygous diploid humanized strain shows no lethality ($CFU_O/CFU_E = \sim1$). PCR-based genotyping of surviving diploid strains after MERGE[1] shows the conversion of both yeast to the humanized loci (ii) compared with before MERGE[1] (i). Sporulation generated the paired-humanized haploids in both mating types (shown here as Tetrads spotted on YPD, iii, and *MATα* 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) Mating each single-humanized strain allows a systematic test for viable double-humanized genotype (21 genotypes). Transformation of double-sgRNA-CRISPR plasmid targeting yeast genes in each paired heterozygote facilitates the combination of two humanized loci as homozygotes ($CFU_O/CFU_E = \sim1$, yellow) while also revealing genotypes that are not permitted ($CFU_O/CFU_E = \sim0$, blue). Data are shown as heatmap and performed in triplicate.

(D) Mixing and mating singly humanized genotypes (*Hsα1*, *Hsα3*, and *Hsα7*) allows a random combination of two heterozygous human-yeast alleles. The transformation of the double-sgRNA-CRISPR plasmid (referred to as mix, mate, and MERGE [MERGE$^{M\&M}$]) in a mated mix enables selection of combined homozygous humanized genotypes. Transformation of pCas9 alone serves as a control (i). Double-sgRNA-CRISPR plasmids, pCas9-sgRNA$^{Sc\text{-}\alpha1,\alpha7}$ (ii) and pCas9-sgRNA$^{Sc\text{-}\alpha3,\alpha7}$ (iii), specifically enriched the corresponding double-humanized genotype while also converting yeast to humanized loci. PCR-based genotyping of randomly picked colonies confirm *Hsα1,α7* and *Hsα3,α7* homozygous genotypes. Comparatively, the transformation of a double-sgRNA-CRISPR plasmid, pCas9-sgRNA$^{Sc\text{-}\alpha1,\alpha6}$ (iv), shows no surviving colonies for a non-existing genotype.

CellPress
OPEN ACCESS

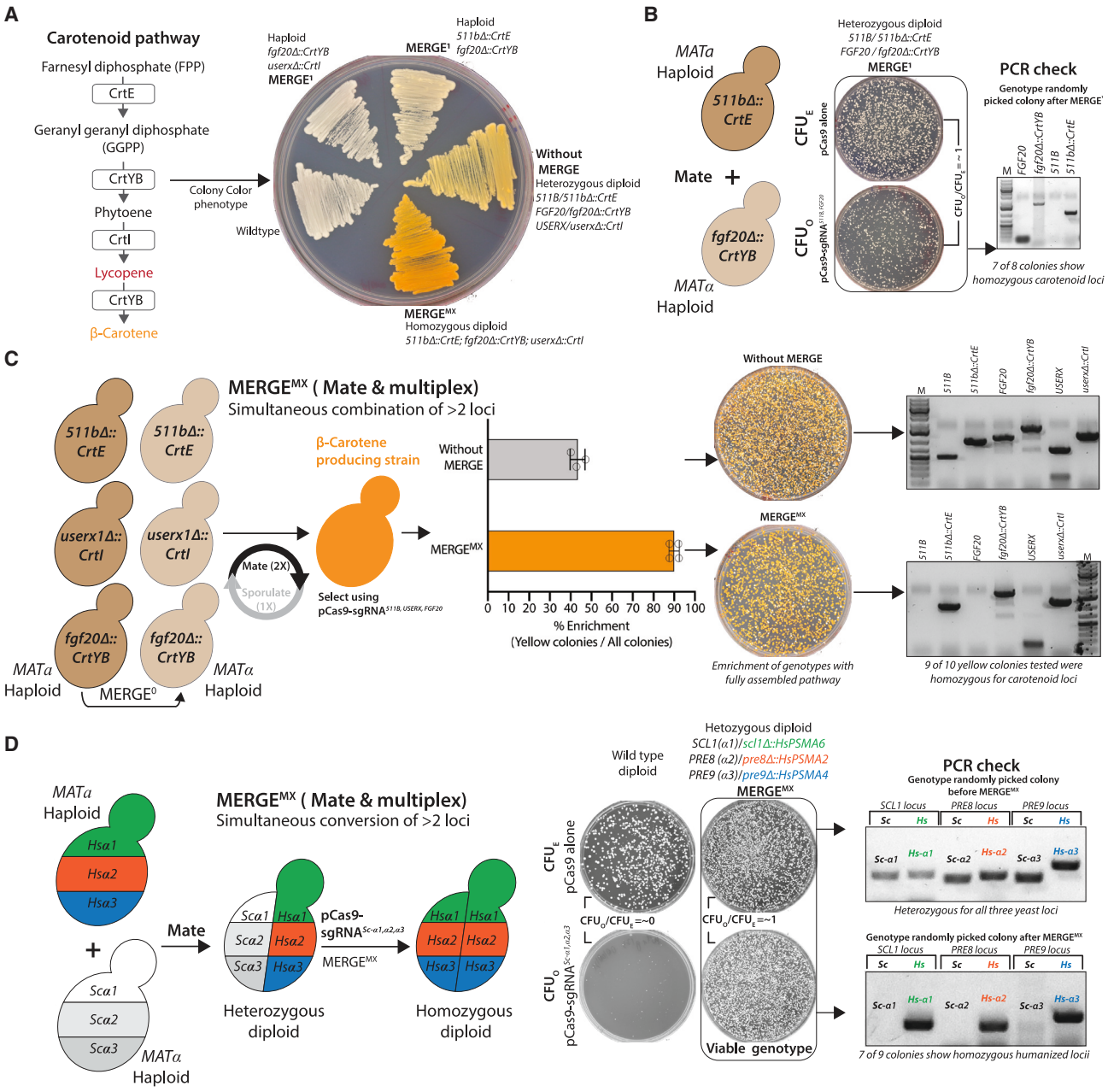


Figure 5. MERGE^{MX} is scalable to combine multiple engineered loci

(A) Schematic shows the carotenoid pathway genes and the metabolic intermediates leading to the color colony phenotype. A complete pathway (*CrtE*, *CrtYB*, and *CrtI*) leads to an orange colony appearance, whereas the partial assembly (*CrtYB* and *CrtI*) provides an off-white colony phenotype. MERGE^{1} provided single-carotenoid engineered strains in both mating types. The homozygous diploid for a complete carotenoid pathway shows a more intense orange colony phenotype than the heterozygous strain.

(B) Schematic showing MERGE^{1} combining two engineered genotypes at landing pad loci. A double-sgRNA-CRISPR plasmid (pCas9-sgRNA^{511B,FGF20}) targeting two landing pad loci (*511B* and *FGF20*) quantifies the efficiency of MERGE^{1}. The transformation of pCas9-sgRNA^{511B,FGF20} in the double heterozygous diploid strain shows no lethality ($CFU_O/CFU_E = \sim1$). PCR-based genotyping of several colonies after MERGE^{1} shows the conversion to the engineered carotenoid loci.

(C) Single-carotenoid gene strains of both mating-types were mixed, mated (2×), and sporulated (1×). The mixture transformed with pCas9 alone shows a distribution of white, off-white, and orange-colored colonies. However, the triple-sgRNA-pCas9-sgRNA^{511B,FGF20,USERX} transformed mix significantly enriched the yellow colonies (data are shown as mean with standard deviation performed in triplicate). PCR-based genotyping of several colonies showed that MERGE^{MX} successfully homozygosed all loci, whereas the yellow colonies on the unselected plate showed heterozygous loci.

(D) Schematics show that mating haploid triple-humanized *Hsα1,α2,α3* strain with a wild-type yeast enables the combination of 3 humanized loci as heterozygotes. Triple-sgRNA-CRISPR plasmid (pCas9-sgRNA$^{Sc\text{-}\alpha1,\alpha2,\alpha3}$) targeted to the corresponding wild-type yeast alleles (CRISPR-sensitive loci) simultaneously initiates recombination at all 3 loci using the homologous chromosome with humanized CRISPR-resistant loci as a repair template. pCas9-sgRNA$^{Sc\text{-}\alpha1,\alpha2,\alpha3}$

(legend continued on next page)

CRISPR selections and exogenous human gene repair templates (Figure 6A). The co-transformation of a triple-sgRNA-CRISPR plasmid (pCas9-sgRNA$^{Sc\text{-}\alpha1,\alpha2,\alpha3}$) targeting yeast *α1*, *α2*, and *α3* genes and PCR fragments of human gene repair templates was successful in obtaining a triple-humanized strain (Figure S7Ai). A similar strategy generated other combinations of triple-humanized strains (Figure S7Ai) but failed to yield a quadruple-humanized strain (Figure S7Aii). Thus, yeast genes are replaceable sequentially, either alone or as small-scale simultaneous replacements. Using a triple-humanized genotype (*Hsα1,α2,α3*) as a background strain, the humanization of yeast *α4* was successful. Next, we explored *Hsα5*, *Hsα6*, and *Hsα7* humanizations in parallel; however, we obtained only one quintuple-humanized (*Hsα1,α2,α3,α4,α7*) genotype (Figures 6A and S7B). The functional replacement of yeast *α5* or *α6* was unsuccessful despite repeated attempts. The plasmid-borne expression *Hsα6* in a quintuple-humanized strain (*Hsα1,α2,α3,α4,α7*) resulted in a toxic phenotype (no growth), suggesting that further humanizations may be incompatible (Figure 6A). In general, while partly successful, the sequential strategy failed to reveal if the inability to fully humanize yeast α-proteasome core was due to incompatible genotypes or inefficient genome editing, especially as hybrid human-yeast genotypes show growth defects and reduced transformation efficiencies (Figure S7C).

In contrast, MERGE^{MX} provided a clear readout of all humanized genotypes tested, readily generating many combinations of humanized α-proteasome genes (Figure 6B). Mating of various humanized strains followed by MERGE^{MX} explored several triple-humanized genotypes, successfully identifying viable (Figures 5C and S6A) and inviable combinations (Figures S6B and S6C). Next, we investigated higher-order (>3) combinations, exploring quadruple- and quintuple-humanized genotypes (Figures S7D–S7G). Finally, MERGE yielded a viable sextuple-humanized (*Hsα1,α2,α3,α4,α6,α7*) genotype with a slow growth phenotype (Figures 6B and S7H). To test whether the entire yeast α-proteasome core is humanizable, we generated a diploid yeast strain harboring all the α-proteasome genes as heterozygous human-yeast genotypes (Figure S8A). MERGE^{MX}, using the triple-sgRNA-CRISPR plasmid, converted yeast *α1*, *α2*, and *α3* loci to homozygous human alleles. However, the subsequent conversion of the remaining four yeast to human loci did not yield viable colonies, suggesting that a fully human α-proteasome core is incompatible (Figure S8B). Overall, the 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, S5D, S5D', and S8C). These data confirm that the yeast proteasome subunits α5 and α6 are not co-humanizable.

The sporulation failure observed in genomically replaced strains can lead to dead ends while performing MERGE. We propose two solutions: one, by allowing the strains with heterozygous engineered loci to sporulate without MERGE, followed by using CRISPR plasmid selection to enrich combined haploid genotypes (Figure S8D), or two, by using a sequential strategy to engineer a viable genotype (Figures S8E). Furthermore, to rule out multiplexed Cas9-sgRNAs-induced off-target effects, we performed whole-genome sequencing of single-humanized (*Hsα1*) and quintuple-humanized strains (*Hsα1α2α3α4α7*), observing no off-target mutations.

Proteasome biogenesis is a highly regulated process aided by several assembly chaperones.[30–32] In the case of α-proteasome core, particularly α5 and α6, subunits interact with assembly chaperones, enabling ordered assembly.[33–35] We performed mass spectrometry (MS) analysis to show proteome-wide changes with significant over-expression of many yeast proteasome subunits, including several assembly chaperones in the humanized proteasome strains (Figure S9A; Table S2). Furthermore, the β-core assembly immediately follows the α-core, and the incompatible interface may now require human β subunits in yeast.[30,36] 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. Indeed, the co-expression of human constitutive- and immuno-βs followed by MS analysis in the singly humanized *Hsα1*, quintuple-humanized *Hsα1,α2,α3,α4,α7,* and sextuple-humanized *Hsα1,α2,α3,α4,α7,β3* (also harboring human β3 in place of the yeast ortholog) strains show stable expression of human β subunits only in partially humanized yeast strains (Figure S9A). Together with the phenotypic rescue of the humanized yeast strains by human β subunits, these data suggest the assembly of a human 20S proteasome core (Figures S9B and S9C), further supported by our recent efforts to humanize the yeast β-core.[37] Given the complex synthesis and architecture of the proteasome,[38] it is challenging to know if there are a limited number of "paths" to humanizing yeast 20S core particle due to a rapidly accumulating number of assays to perform. However, using MERGE together with co-expression of critical human proteasome-associated subunits (as in Figure S9) and mutational profiling of non-replaceable human genes,[26,37,39] yeast with an entirely humanized catalytically active 20S proteasome core is possible.

DISCUSSION

Yeast genome engineering has entered a new era, with CRISPR-Cas9 enabling multiplexed strategies and gene drives to edit several loci simultaneously.[22,40] The advantages of using CRISPR selection are numerous, with no dependency on markers for any number of genotypes. By integrating marker-less CRISPR-Cas9-based selection with cycles of mating and sporulation, we developed a method (MERGE) that enables a highly efficient combination of distinct loci in yeast. Furthermore, given the efficiency of MERGE and because the method requires the creation of viable heterozygous intermediate strains, it permits the quantitative identification of all possible combined genotypes using a simple readout of CFUs.

transformation in the wild type is lethal ($CFU_O/CFU_E = \sim0$); however, the diploid triple heterozygous humanized strain shows no lethality ($CFU_O/CFU_E = \sim1$). PCR-based genotyping of surviving colonies after MERGE^{MX} confirms the conversion of all 3 yeast to the humanized loci compared with before MERGE^{MX}. The transformation of a plasmid without sgRNA (pCas9 alone) serves as a transformation efficiency control (CFU_E).

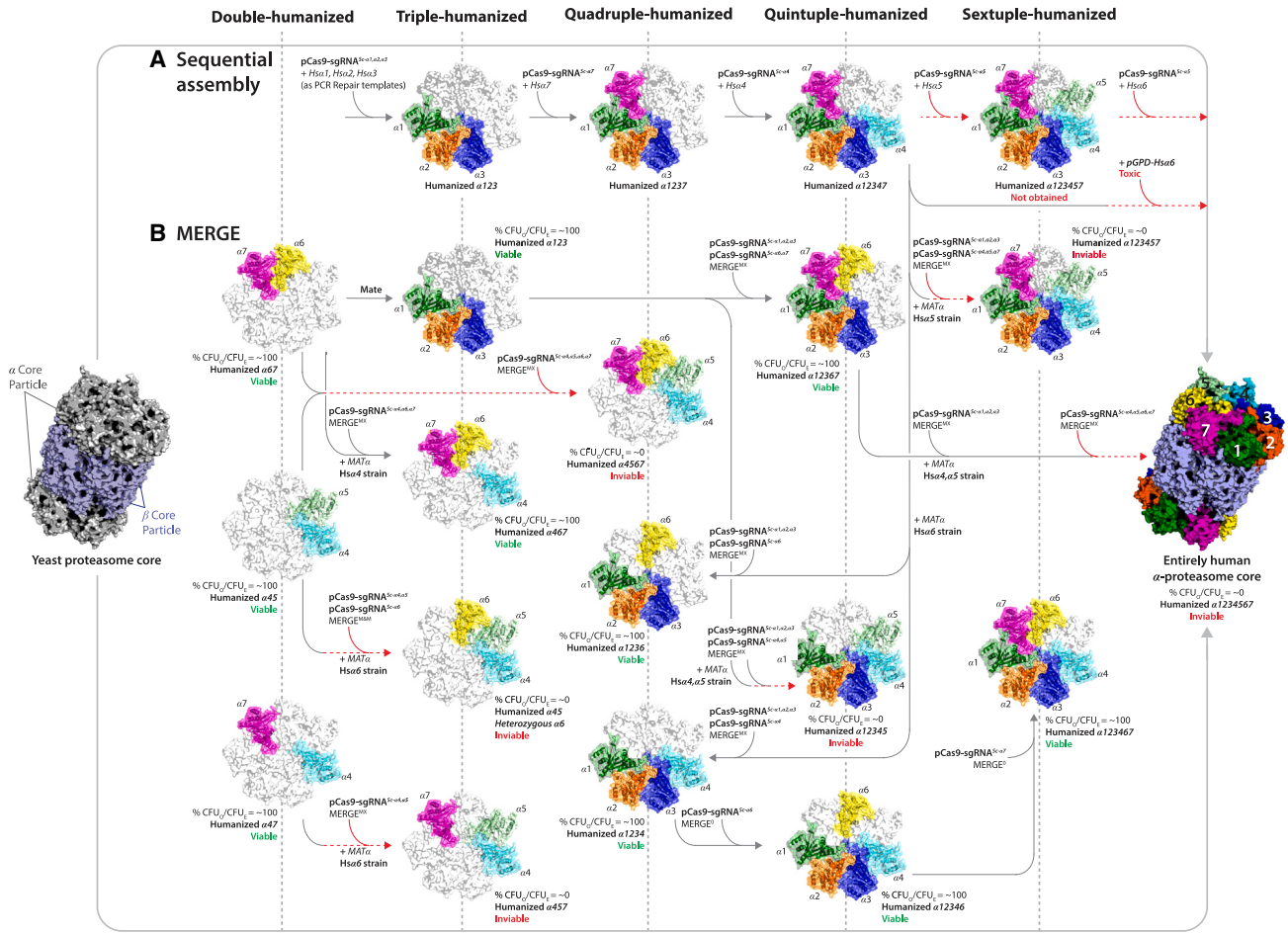


Figure 6. MERGE reveals a fitness-driven path to the humanization of most of the α-proteasome core in yeast

The schematic shows the transition of heptameric yeast α-proteasome to humanized α-proteasome core.

(A) A sequential strategy used a triple-humanized *Hsα1,α2,α3* strain as a background to progressively humanize the rest of yeast α-proteasome core genes by co-transforming a CRISPR plasmid and a human gene repair template. Choices to sequentially humanize the yeast α-proteasome core genes were made depending on the success of the prior effort. The strategy permitted the humanization of yeast *α7* followed by *α4*, whereas several attempts to humanize yeasts *α5* and *α6* failed. The plasmid-based expression of human Hsα6 in a quintuple-humanized *Hsα1,α2,α3,α4,α7* strain resulted in a lethal phenotype.

(B) MERGE provided a clear readout of the fitness of combined genotypes while revealing incompatible combinations of the humanized yeast α-proteasome core. Mating distinct humanized yeast combinations (shown as connecting lines) followed by MERGE tested various triple-, quadruple-, quintuple-, and sextuple-humanized α-proteasome genes. MERGE generated viable triple-humanized genotypes while identifying incompatible combinations (indicated as dashed lines in red). Similarly, while MERGE^{1} facilitated obtaining *Hsα4,α5* and *Hsα6,α7* genotypes, the subsequent MERGE^{MX} to generate quadrupled *Hsα4,α5,α6,α7* failed. Using MERGE^{MX} followed by MERGE^{0} identified several quadruple- and quintuple- and one sextuple-humanized genotype (*Hsα1,α2,α3,α4,α6,α7*), revealing a fitness-driven path to the humanization of the near-entire yeast α-proteasome core. Several humanized combinations (>2), comprising the Hsα5 subunit, were incompatible. Mating of *Hsα1,α2,α3,α6,α7* and *Hsα4,α5* strains generated a heterozygous human-yeast diploid for all 7 α-proteasome core genes. Using multiplex-sgRNA-CRISPR plasmids permitted verifying if the entire human α-proteasome core is compatible in yeast. A triple-sgRNA-CRISPR plasmid (pCas9-sgRNA$^{Sc\text{-}\alpha1,\alpha2,\alpha3}$) homozygosed 3 of 7 humanized loci. However, a subsequent MERGE^{MX}, using a quadruple-sgRNA-CRISPR plasmid (pCas9-sgRNA$^{Sc\text{-}\alpha4,\alpha5,\alpha6,\alpha7}$), failed to obtain a viable genotype. Proteasome core structures were generated using Pymol and PDB-1RYP. Colored structures show the humanized α-proteasome subunits.

MERGE offers highly scalable multi-locus genome engineering in diploid yeast cells using high-efficiency CRISPR-based gene drives to overcome the independent assortment of unlinked loci. While the principle is similar, the variations of the method may be used to explore various functional genomics questions, for example, MERGE^{0} for haploinsufficiency or essential gene assays,[41] MERGE^{1} for synthetic genetic interactions,[42] MERGE^{MX} for multiplex editing and higher-order genetic interactions (>3),[43] and MERGE$^{M\&M}$ for evolutionary experiments to select and enrich genotypes from a mixture. In the future, MERGE could potentially drive synthetic and systems biology research from assembling heterologous systems to performing multi-site and genome-wide combinatorial editing, as we demonstrated by editing 89 independent sites along chromosome 1; 16 *GFP* insertions within

the same strain, engineering a complete four-gene carotenoid biosynthesis pathway; and 37 single- or multiple-humanized proteasome genes in yeast.

Given the broad conservation of DNA repair machinery, MERGE can simplify systematic functional genomic analysis in model systems, such as *Candida* and fission yeast. The double-strand break repair via HDR pathway is highly conserved in higher eukaryotes, including humans.[44] Therefore, MERGE may be similarly effective in higher eukaryotes. CRISPR plasmids that exclusively target one of the alleles of a heterozygous human locus can quantify MERGE efficiency. If successful, the strategy could potentially correct human disease-associated heterozygous alleles at a significantly higher rate than conventional methods using exogenous repair templates.

By humanizing 6 of 7 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 is driven by modules of physically or functionally interacting proteins being similarly replaceable.[26] Further characterization of the incompatibilities should reveal 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 an unexplored 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.

Limitations of the study

Our method combines individual genetic edits using haploid yeast, mating, and CRISPR-Cas9-induced gene drives. The strategy requires that each edit generate a viable yeast strain first. Therefore, the method is not applicable if the yeast strains with individual edits are inviable, as with many functionally non-replaceable human genes.[26,45] However, using a diploid strain to generate heterozygous alleles will enable combining the edits using MERGE. Furthermore, we explore MERGE in yeast that uses inherently efficient HR to repair DSBs. Thus, MERGE-based strategies will require highly effective HR in other model organisms.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.crmeth.2023.100464.

ACKNOWLEDGMENTS

The authors would like to thank Drs. Charles Boone (University of Toronto) for sharing the SGA-compatible yeast strains, Frederick Roth (University of Toronto) and Harvard University for sharing GM 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. This research was funded by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) (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 was from School of Graduate Studies (SGS), Faculty of Arts and Sciences, Concordia University to M.A. and B.M.G. and the 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).

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

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: September 2, 2022
Revised: January 30, 2023
Accepted: April 12, 2023
Published: May 10, 2023

REFERENCES

1. Botstein, D., and Fink, G.R. (2011). Yeast: an experimental organism for 21st Century biology. Genetics *189*, 695–704. https://doi.org/10.1534/genetics.111.130765.
2. Costanzo, M., VanderSluis, B., Koch, E.N., Baryshnikova, A., Pons, C., Tan, G., Wang, W., Usaj, M., Hanchard, J., Lee, S.D., et al. (2016). A global genetic interaction network maps a wiring diagram of cellular function. Science *353*, aaf1420. https://doi.org/10.1126/science.aaf1420.
3. Richardson, S.M., Mitchell, L.A., Stracquadanio, G., Yang, K., Dymond, J.S., DiCarlo, J.E., Lee, D., Huang, C.L.V., Chandrasegaran, S., Cai, Y., et al. (2017). Design of a synthetic yeast genome. Science *355*, 1040–1044. https://doi.org/10.1126/science.aaf4557.
4. Suzuki, Y., Assad-Garcia, N., Kostylev, M., Noskov, V.N., Wise, K.S., Karas, B.J., Stam, J., Montague, M.G., Hanly, T.J., Enriquez, N.J., et al. (2015). Bacterial genome reduction using the progressive clustering of deletions via yeast sexual cycling. Genome Res. *25*, 435–444. https://doi.org/10.1101/gr.182477.114.
5. Barbieri, E.M., Muir, P., Akhuetie-Oni, B.O., Yellman, C.M., and Isaacs, F.J. (2017). Precise editing at DNA replication forks enables multiplex genome engineering in eukaryotes. Cell *171*, 1453–1467.e13. https://doi.org/10.1016/j.cell.2017.10.034.
6. Gibson, D.G., Benders, G.A., Andrews-Pfannkoch, C., Denisova, E.A., Baden-Tillson, H., Zaveri, J., Stockwell, T.B., Brownley, A., Thomas, D.W., Algire, M.A., et al. (2008). Complete chemical synthesis, assembly, and cloning of a Mycoplasma genitalium genome. Science *319*, 1215–1220. https://doi.org/10.1126/science.1151721.
7. Fredens, J., Wang, K., de la Torre, D., Funke, L.F.H., Robertson, W.E., Christova, Y., Chia, T., Schmied, W.H., Dunkelmann, D.L., Beránek, V., et al. (2019). Total synthesis of Escherichia coli with a recoded genome. Nature *569*, 514–518. https://doi.org/10.1038/s41586-019-1192-5.
8. Muller, H., Annaluru, N., Schwerzmann, J.W., Richardson, S.M., Dymond, J.S., Cooper, E.M., Bader, J.S., Boeke, J.D., and Chandrasegaran, S. (2012). Assembling large DNA segments in yeast. Methods Mol. Biol. *852*, 133–150. https://doi.org/10.1007/978-1-61779-564-0_11.
9. Annaluru, N., Muller, H., Mitchell, L.A., Ramalingam, S., Stracquadanio, G., Richardson, S.M., Dymond, J.S., Kuang, Z., Scheifele, L.Z., Cooper, E.M., et al. (2014). Total synthesis of a functional designer eukaryotic chromosome. Science *344*, 55–58. https://doi.org/10.1126/science.1249252.
10. Lajoie, M.J., Rovner, A.J., Goodman, D.B., Aerni, H.-R., Haimovich, A.D., Kuznetsov, G., Mercer, J.A., Wang, H.H., Carr, P.A., Mosberg, J.A., et al. (2013). Genomically recoded organisms expand biological functions. Science *342*, 357–360. https://doi.org/10.1126/science.1241459.
11. Truong, D.M., and Boeke, J.D. (2017). Resetting the yeast epigenome with human nucleosomes. Cell *171*, 1508–1519.e13. https://doi.org/10.1016/j.cell.2017.10.043.
12. Agmon, N., Temple, J., Tang, Z., Schraink, T., Baron, M., Chen, J., Mita, P., Martin, J.A., Tu, B.P., Yanai, I., et al. (2020). Phylogenetic debugging of a complete human biosynthetic pathway transplanted into yeast. Nucleic Acids Res. *48*, 486–499. https://doi.org/10.1093/nar/gkz1098.
13. Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281–2308. https://doi.org/10.1038/nprot.2013.143.
14. DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., and Church, G.M. (2013). Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res. *41*, 4336–4343. https://doi.org/10.1093/nar/gkt135.
15. Akhmetov, A., Laurent, J.M., Gollihar, J., Gardner, E.C., Garge, R.K., Ellington, A.D., Kachroo, A.H., and Marcotte, E.M. (2018). Single-step precision genome editing in yeast using CRISPR-Cas9. Bio. Protoc. *8*, e2765. https://doi.org/10.21769/BioProtoc.2765.
16. Kachroo, A.H., Laurent, J.M., Akhmetov, A., Szilagyi-Jones, M., McWhite, C.D., Zhao, A., and Marcotte, E.M. (2017). Systematic bacterialization of yeast genes identifies a near-universally swappable pathway. Elife *6*, e25093. https://doi.org/10.7554/eLife.25093.
17. Lee, M.E., DeLoache, W.C., Cervantes, B., and Dueber, J.E. (2015). A highly characterized yeast toolkit for modular, multipart assembly. ACS Synth. Biol. *4*, 975–986. https://doi.org/10.1021/sb500366v.
18. Kuzmin, E., Costanzo, M., Andrews, B., and Boone, C. (2016). Synthetic genetic array analysis. Cold Spring Harb. Protoc. *2016*, pdb.prot088807. https://doi.org/10.1101/pdb.prot088807.
19. Suzuki, Y., St Onge, R.P., Mani, R., King, O.D., Heilbut, A., Labunskyy, V.M., Chen, W., Pham, L., Zhang, L.V., Tong, A.H.Y., et al. (2011). Knocking out multigene redundancies via cycles of sexual assortment and fluorescence selection. Nat. Methods *8*, 159–164. https://doi.org/10.1038/nmeth.1550.
20. Labunskyy, V.M., Suzuki, Y., Hanly, T.J., Murao, A., Roth, F.P., and Gladyshev, V.N. (2014). The insertion Green Monster (iGM) method for expression of multiple exogenous genes in yeast. G3 *4*, 1183–1191. https://doi.org/10.1534/g3.114.010868.
21. Xu, H., Han, M., Zhou, S., Li, B.-Z., Wu, Y., and Yuan, Y.-J. (2020). Chromosome drives via CRISPR-Cas9 in yeast. Nat. Commun. *11*, 4344. https://doi.org/10.1038/s41467-020-18222-0.
22. Yan, Y., and Finnigan, G.C. (2018). Development of a multi-locus CRISPR gene drive system in budding yeast. Sci. Rep. *8*, 17277. https://doi.org/10.1038/s41598-018-34909-3.
23. DiCarlo, J.E., Chavez, A., Dietz, S.L., Esvelt, K.M., and Church, G.M. (2015). Safeguarding CRISPR-Cas9 gene drives in yeast. Nat. Biotechnol. *33*, 1250–1255. https://doi.org/10.1038/nbt.3412.
24. Roggenkamp, E., Giersch, R.M., Wedeman, E., Eaton, M., Turnquist, E., Schrock, M.N., Alkotami, L., Jirakittisonthon, T., Schluter-Pascua, S.E., Bayne, G.H., et al. (2017). CRISPR-UnLOCK: multipurpose Cas9-based strategies for conversion of yeast libraries and strains. Front. Microbiol. *8*, 1773. https://doi.org/10.3389/fmicb.2017.01773.
25. Kramer, K.M., Brock, J.A., Bloom, K., Moore, J.K., and Haber, J.E. (1994). Two different types of double-strand breaks in Saccharomyces cerevisiae are repaired by similar RAD52-independent, nonhomologous recombination events. Mol. Cell Biol. *14*, 1293–1301. https://doi.org/10.1128/mcb.14.2.1293-1301.1994.
26. Kachroo, A.H., Laurent, J.M., Yellman, C.M., Meyer, A.G., Wilke, C.O., and Marcotte, E.M. (2015). Evolution. Systematic humanization of yeast genes reveals conserved functions and genetic modularity. Science *348*, 921–925. https://doi.org/10.1126/science.aaa0769.

27. Winzeler, E.A., Shoemaker, D.D., Astromoff, A., Liang, H., Anderson, K., Andre, B., Bangham, R., Benito, R., Boeke, J.D., Bussey, H., et al. (1999). Functional characterization of the S. cerevisiae genome by gene deletion and parallel analysis. Science *285*, 901–906.

28. Pan, X., Yuan, D.S., Xiang, D., Wang, X., Sookhai-Mahadeo, S., Bader, J.S., Hieter, P., Spencer, F., and Boeke, J.D. (2004). A robust toolkit for functional profiling of the yeast genome. Mol. Cell *16*, 487–496.

29. Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823. https://doi.org/10.1126/science.1231143.

30. Budenholzer, L., Cheng, C.L., Li, Y., and Hochstrasser, M. (2017). Proteasome structure and assembly. J. Mol. Biol. *429*, 3500–3524. https://doi.org/10.1016/j.jmb.2017.05.027.

31. Hirano, Y., Hayashi, H., Iemura, S.-I., Hendil, K.B., Niwa, S.-I., Kishimoto, T., Kasahara, M., Natsume, T., Tanaka, K., and Murata, S. (2006). Cooperation of multiple chaperones required for the assembly of mammalian 20S proteasomes. Mol. Cell *24*, 977–984. https://doi.org/10.1016/j.molcel.2006.11.015.

32. Le Tallec, B., Barrault, M.-B., Courbeyrette, R., Guérois, R., Marsolier-Kergoat, M.-C., and Peyroche, A. (2007). 20S proteasome assembly is orchestrated by two distinct pairs of chaperones in yeast and in mammals. Mol. Cell *27*, 660–674. https://doi.org/10.1016/j.molcel.2007.06.025.

33. Schnell, H.M., Walsh, R.M., Rawson, S., Kaur, M., Bhanu, M.K., Tian, G., Prado, M.A., Guerra-Moreno, A., Paulo, J.A., Gygi, S.P., et al. (2021). Structures of chaperone-associated assembly intermediates reveal coordinated mechanisms of proteasome biogenesis. Nat. Struct. Mol. Biol. *28*, 418–425. https://doi.org/10.1038/s41594-021-00583-9.

34. Morris, E.P., and da Fonseca, P.C.A. (2021). How to build a proteasome. Nat. Struct. Mol. Biol. *28*, 409–410. https://doi.org/10.1038/s41594-021-00592-8.

35. Stadtmueller, B.M., Kish-Trier, E., Ferrell, K., Petersen, C.N., Robinson, H., Myszka, D.G., Eckert, D.M., Formosa, T., and Hill, C.P. (2012). Structure of a proteasome Pba1-Pba2 complex: implications for proteasome assembly, activation, and biological function. J. Biol. Chem. *287*, 37371–37382. https://doi.org/10.1074/jbc.M112.367003.

36. Murata, S., Yashiroda, H., and Tanaka, K. (2009). Molecular mechanisms of proteasome assembly. Nat. Rev. Mol. Cell Biol. *10*, 104–115. https://doi.org/10.1038/nrm2630.

37. Sultana, S., Abdullah, M., Li, J., Hochstrasser, M., and Kachroo, A.H. (2022). Species-specific protein-protein interactions govern the humanization of the 20S proteasome in yeast. Preprint at bioRxiv. https://doi.org/10.1101/2022.06.20.496808.

38. Tomko, R.J., and Hochstrasser, M. (2013). Molecular architecture and assembly of the eukaryotic proteasome. Annu. Rev. Biochem. *82*, 415–445. https://doi.org/10.1146/annurev-biochem-060410-150257.

39. Kachroo, A.H., Vandeloo, M., Greco, B.M., and Abdullah, M. (2022). Humanized yeast to model human biology, disease and evolution. Dis. Model. Mech. *15*, dmm049309. https://doi.org/10.1242/dmm.049309.

40. Ryan, O.W., Skerker, J.M., Maurer, M.J., Li, X., Tsai, J.C., Poddar, S., Lee, M.E., DeLoache, W., Dueber, J.E., Arkin, A.P., et al. (2014). Selection of chromosomal DNA libraries using a multiplex CRISPR system. Elife *3*, e03703. https://doi.org/10.7554/eLife.03703.

41. Smith, A.M., Ammar, R., Nislow, C., and Giaever, G. (2010). A survey of yeast genomic assays for drug and target discovery. Pharmacol. Ther. *127*, 156–164. https://doi.org/10.1016/j.pharmthera.2010.04.012.

42. Tong, A.H., Evangelista, M., Parsons, A.B., Xu, H., Bader, G.D., Pagé, N., Robinson, M., Raghibizadeh, S., Hogue, C.W., Bussey, H., et al. (2001). Systematic genetic analysis with ordered arrays of yeast deletion mutants. Science *294*, 2364–2368.

43. Kuzmin, E., Andrews, B.J., and Boone, C. (2021). Trigenic synthetic genetic array (τ-SGA) technique for complex interaction analysis. Methods Mol. Biol. *2212*, 377–400. https://doi.org/10.1007/978-1-0716-0947-7_23.

44. Kanaar, R., Troelstra, C., Swagemakers, S.M., Essers, J., Smit, B., Fransen, J.H., Pastink, A., Bezzubova, O.Y., Buerstedde, J.M., Clever, B., et al. (1996). Human and mouse homologs of the Saccharomyces cerevisiae RAD54 DNA repair gene: evidence for functional conservation. Curr. Biol. *6*, 828–838. https://doi.org/10.1016/s0960-9822(02)00606-1.

45. Laurent, J.M., Garge, R.K., Teufel, A.I., Wilke, C.O., Kachroo, A.H., and Marcotte, E.M. (2020). Humanization of yeast genes with multiple human orthologs reveals functional divergence between paralogs. PLoS Biol. *18*, e3000627. https://doi.org/10.1371/journal.pbio.3000627.

46. Tong, A.H.Y., and Boone, C. (2006). Synthetic genetic array analysis in Saccharomyces cerevisiae. Methods Mol. Biol. *313*, 171–192. https://doi.org/10.1385/1-59259-958-3:171.

47. Mitchell, L.A., Chuang, J., Agmon, N., Khunsriraksakul, C., Phillips, N.A., Cai, Y., Truong, D.M., Veerakumar, A., Wang, Y., Mayorga, M., et al. (2015). Versatile genetic assembly system (VEGAS) to assemble pathways for expression in S. cerevisiae. Nucleic Acids Res. *43*, 6620–6630. https://doi.org/10.1093/nar/gkv466.

48. Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics *28*, 1647–1649. https://doi.org/10.1093/bioinformatics/bts199.

49. Bourgeois, L., Pyne, M.E., and Martin, V.J.J. (2018). A highly characterized synthetic landing pad system for precise multicopy gene integration in yeast. ACS Synth. Biol. *7*, 2675–2685. https://doi.org/10.1021/acssynbio.8b00339.

50. Lamesch, P., Li, N., Milstein, S., Fan, C., Hao, T., Szabo, G., Hu, Z., Venkatesan, K., Bethel, G., Martin, P., et al. (2007). hORFeome v3.1: a resource of human open reading frames representing over 10,000 human genes. Genomics *89*, 307–315. https://doi.org/10.1016/j.ygeno.2006.11.012.

51. MGC Project Team; Temple, G., Gerhard, D.S., Rasooly, R., Feingold, E.A., Good, P.J., Robinson, C., Mandich, A., Derge, J.G., Lewis, J., et al. (2009). The completion of the mammalian gene collection (MGC). Genome Res. *19*, 2324–2333.

52. Lõoke, M., Kristjuhan, K., and Kristjuhan, A. (2011). Extraction of genomic DNA from yeasts for PCR-based applications. Biotechniques *50*, 325–328. https://doi.org/10.2144/000113672.

53. Gietz, R.D., and Woods, R.A. (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Meth. Enzymol. *350*, 87–96.

54. Huh, W.-K., Falvo, J.V., Gerke, L.C., Carroll, A.S., Howson, R.W., Weissman, J.S., and O'Shea, E.K. (2003). Global analysis of protein localization in budding yeast. Nature *425*, 686–691. https://doi.org/10.1038/nature02026.

55. Suzuki, Y., Stam, J., Novotny, M., Yachie, N., Lasken, R.S., and Roth, F.P. (2012). The green monster process for the generation of yeast strains carrying multiple gene deletions. J. Vis. Exp., e4072. https://doi.org/10.3791/4072.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, Peptides, and Recombinant Proteins		
YPD Broth (Powder)	Sigma-Aldrich	Cat#Y1375
YPD Agar	Sigma-Aldrich	Cat#Y1500-
YNB powder	Sigma-Aldrich	Cat# Y1251
Bacto-agar	Sigma-Aldrich	Cat#A5306-
Dextrose	Sigma-Aldrich	Cat# G8270
LB Agar	Sigma-Aldrich	Cat#L2897
LB broth	Sigma-Aldrich	Cat#L3022-
Drop-out Medium without Uracil	Sigma-Aldrich	Cat#Y1501
Drop-out Medium without Histidine	Sigma-Aldrich	Cat#Y1751
Drop-out Medium without Leucine	Sigma-Aldrich	Cat#Y1376
5-Fluoro Orotic Acid Monohydrate (5-FOA)	Thermo Fisher Scientific	Cat#R0811
Molecular Grade Water	VWR	Cat#46-000-CI
Zymo Yeast transformation Kit	VWR	Cat#T2002/T2003
SYBR Safe DNA Gel Stain	Thermo Fisher Scientific	Cat# S33102
TAE Buffer 50x	VWR	Cat#97063-696
Accuprime polymerase	Thermo Fisher Scientific	Cat#12-344-024
2X Plant Phire Master Mix	Thermo Fisher Scientific	Cat# F125S
BsaI	Thermo Fisher Scientific	Cat#FD0293
BsmbI	Thermo Fisher Scientific	Cat#FD0454
ATP solution (10mM)	Thermo Fisher Scientific	Cat#AM8110G
T7 ligase	NEB	Cat#M0318S
T4 ligase	NEB	Cat#M0202S
R-Zymolase	Zymoresearch	Cat#E1006
G418	Thermo Fisher Scientific	Cat#LS10131027
Zymolase, 100T9 (10 mg/mL)	(US BIOLOGICAL)	Cat#Z1005
Canavanine	Sigma-Aldrich	Cat#C9758-1G
Carbenicillin	Sigma-Aldrich	Cat#C1389-1g
Molecular Grade Water	Thermo Fisher Scientific	Cat#46-000-CI
L-4-Thialysine hydrochloride	Sigma-Aldrich	Cat#A2636
Pierce™ TCEP-HCl	ThermoFisher	Cat# 20490
Experimental models: Organisms/strains		
Saccharomyces cerevisiae BY4741 (*MATa*, *his3Δ1, leu2Δ0, met15Δ0, ura3Δ0)*	Saccharomyces Genome Database	N/A
Saccharomyces cerevisiae SGA strains (*MATa,can1Δ::STE2pr_HIS3,his3Δ1, leu2Δ0,ura3Δ0)* and *MATα,lyp1Δ::STE3pr_ LEU2,his3Δ1,leu2Δ0,ura3Δ0*	Tong et al.[46]	N/A
Saccharomyces cerevisiae Green Monster strain (*MATα;lyp1Δ,his3Δ1,leu2Δ0,ura3Δ0, met15Δ0,can1Δ:: GMToolkit-α [CMVpr-rtTA NatMX4, STE3pr_LEU2)*	Suzuki et al.[19]	N/A
NEB high efficiency *E. coli* competent cells	New England Biolabs	Cat#C2987H
Oligonucleotides		
See Table S1	This paper	N/A
Recombinant DNA		
MoClo-YTK	Lee et al.[17]	Addgene; Kit #1000000061

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pCas9-*GFP*do	This Paper	N/A
pCas9-*GFP*^{MX}	This Paper	N/A
pCas9-sgRNA$^{yeast\ locus}$ plasmids	This Paper	N/A
Carotenoid expression vectors	Mitchell et al.[47]	Boeke lab.
Software and algorithms		
Geneious	https://www.geneious.com	N/A
ImageJ	https://ImageJ.nih.gov/ij/download.html	N/A
Pymol	https://pymol.org/2/	N/A
ChimeraX	https://www.cgl.ucsf.edu/chimerax/index.html	N/A
GraphPad Prism version 9.0.1	https://www.graphpad.com	N/A
Excel Office	Microsoft, WA, USA	https://www.microsoft.com/
Adobe Illustrator 2022	Adobe Systems, CA, USA	https://adobe.com/products/illustrator
Adobe Photoshop 2022	Adobe Systems, CA, USA	https://adobe.com/products/illustrator
MORPHEUS	Broad Institute, MIT, MA, USA	https://software.broadinstitute.org/morpheus/
Proteome discoverer 2.2	ThermoFisher	N/A
Other		
BioTek Synergy H1	Agilent	https://www.agilent.com
Sunrise Tecan microplate reader	Tecan	https://shop.tecan.com
Counter (for colonies)	VWR	Cat#23609-102
Aluminum Foil	VWR	Cat#60941-126
Breathable film	VWR	Cat#60941-086
96 well plates (flat bottom; sterile) (Corning)	Fischer Scientific	Cat# #07-200-89
96-well/2 mL Polypropylene Deep Well Plate (Corning)	Fischer Scientific	Cat#3960
Acclaim™ PepMap™ 100 C18 HPLC Columns	ThermoFisher	Cat# 164946
Hyper-Sep C18 SpinTips	ThermoFisher	Cat# 60109-412

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact Aashiq Kachroo (aashiq.kachroo@concordia.ca).

Materials availability

The humanized yeast strains, CRISPR plasmids, and human-gene yeast expression vectors will be made available by lead contact, Aashiq Kachroo (aashiq.kachroo@concordia.ca), upon request.

Data and code availability

- All data reported in this paper will be shared by the lead contact upon request.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact (aashiq.kachroo@concordia.ca), upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Media and culture

Yeast strains were grown in rich Yeast extract Peptone Dextrose (YPD) or Synthetic Defined (SD− [minus] lacking key metabolites for auxotrophic selection) liquid/solid agar media. In addition, SD− His+Canavanine and SD− Leu+Thialysine media was used to select

haploid yeast cells from the mixture of haploids and diploids. Strains harboring *kanMX* cassette were selected using G418 (Geneticin). For specific details, see the "method details".

Saccharomyces cerevisiae strain generation

Yeast transformations were performed using the Frozen-EZ Yeast Transformation II Kit from Zymo Research (Cedarlane) according to the manufacturer's protocol. In general, we transformed approximately 1μg of the CRISPR-sgRNA plasmid and 5μg of repair DNA (human or carotenoid genes), which yielded anywhere from single to dozens of colonies. Nearly all screened colonies showed successful repair. In addition, we tested two sgRNA sequences for each yeast locus and verified ON-target efficiency in the absence of repair template. Successful Cas9-sgRNA shows zero or very few colonies in the absence of a repair DNA template.

Humanized yeast strains were generated in BY4741 *MATα* strains using Cas9-sgRNA plasmids targeting the corresponding yeast loci. Carotenoid gene transcription units harboring strong constitutive yeast promoters were inserted at the landing pad loci (*CrtE* inserted at *511B* locus, CrtI at *USERX*, and CrtYB at *FGF20*) in the SGA (*MATα* and MATα) strains.

The sextuple-humanized *Hsα1,α2,α3,α4,α7,β3* yeast strain was constructed by genomically inserting *Hsβ3* using CRISPR-Cas9-sgRNA strategy in a *Hsα1,α2,α3,α4,α7* strain.

For specific details, see the "method details".

METHOD DETAILS

sgRNA design

CRISPR-Cas9 targeting sequences consist of a 20 bp recognition sequence preceding an 'NGG' sequence motif (PAM). To design sgRNA targeting sequences, we used the built-in guide RNA design tool in version 11 of the Geneious software using a recent version of the yeast genome (available from http://www.yeastgenome.org/strain/S288C/overview) to screen for possible off-target sequences.[15,16,48] We chose two high-scoring guide sequences to target each gene, requiring that they be near the 5′ end of the gene so any NHEJ repair would likely result in an early frameshift. The guides were ordered as complementary oligos with overhangs according to the sgRNA template described in the Yeast Tool Kit[17] (Table S1).

Cas9-single sgRNA plasmid construction using Yeast Tool Kit (YTK)

Entry vectors for each guide RNA sequence were constructed by the BsmbI-mediated Golden Gate (GG) assembly (Thermo) into plasmid pYTK050 from the Yeast ToolKit.[17] The targeting sequences were ordered (IDT) as two single-stranded DNA oligos with a complementary region and unique overhangs according to the 'sgRNA' template described in Lee, ME et al.[17] Briefly, the two oligos for each guide were annealed by slow cooling from 95°C to 4°C (1–5 h) in a thermal cycler, and 1μL of the annealed product was added to the entry vector GG reaction. Entry Vectors were sequence-verified using custom primers.'Transcriptional unit' (TU) vectors (pYTK095) were constructed by a BsaI-mediated GG assembly (Thermo). The entry vector for a particular guide RNA was combined with a left connector part plasmid (ConLX) and right connector part plasmid (ConRX) into the AmpR-ColE1 (pYTK095) backbone plasmid.[15,16] To create single-sgRNA TU vectors, we assembled the sgRNA entry vector with left connector 1 (ConL1, pYTK003) and right connector E (ConRE, pYTK072). To create TU vectors for the multi-sgRNA knockout, an entry vector was combined with the appropriate connectors to enable proper knockout vector assembly (i.e. a triple-sgRNA vector would have one TU vector with ConL1 and R2, two with ConL2 and R3, and the third with ConL3 and RE). 'Knockout' (KO) vectors were constructed by BsmbI-mediated GG assembly (Thermo). The appropriate TU vectors were assembled into the CEN6-URA3 backbone and the Cas9-TU1 vector to create a KO vector. Each GG reaction was performed in a 10μL volume, with approximately 20 fmol of each starting DNA molecule, along with 0.5μL each of the appropriate restriction enzyme, 0.5μL of T7 DNA ligase (NEB), and 1μL of 10X T4 ligase buffer (NEB). A thermocycler was used to cycle between 16°C and 37°C each for (5 min). 3μL of the reaction volume was transformed into competent DH5α *E. coli* and plated on appropriate selective media for the backbone (i.e. chloramphenicol for entry vectors, ampicillin for TU vectors, and kanamycin for KO vectors). Each backbone plasmid is a '*GFP*-dropout' vector, so correct clones were selected by screening for non-fluorescent colonies when viewed by blue light and verified by sequencing.

Cas9-TU1

The Cas9-TU1 was constructed by BsaI GG assembly of YTK parts pYTK002 (ConLS), 011 (PGK1 promoter), 036 (Cas9 coding sequence from *Streptococcus pyogenes*), 055 (ENO2 terminator, 067 (ConR1), and 095 (AmpR-ColE1 *GFP* dropout backbone).

CEN6-URA3 expression vector

The CEN6-URA *GFP* dropout backbone vector was constructed by BsaI GG assembly of YTK parts pYTK008 (ConLS'), pYTK047 (*GFP* dropout), pYTK073 (ConLE'), pYTK074 (URA3), pYTK081 (CEN6/ARS4), pYTK084 (KanR-ColE1 *RFP* dropout vector).

Direct cloning of sgRNAs in a Cas9-yeast expression vector

To perform faster direct sgRNA-Cas9 plasmid construction without requiring three independent cloning strategies, we constructed **pCas9-*GFP*do**. First, the sgRNA-GFP expression unit from the pYTK050 plasmid was amplified via PCR using primers with BsaI sites that generated overhangs similar to ConL1 (forward primer) and ConRE (reverse primer). Next, using the Golden Gate (GG) reaction

protocol, the sgRNA TU PCR was assembled into the CEN6-*URA3* or CEN6-*kanMX* backbone with the Cas9-TU1 vector to create a direct knockout (**pCas9-*GFP*do**) vector.

Direct cloning of multiple sgRNAs transcription units in a Cas9-yeast expression vector

To build >1 sgRNAs TUs directly into the yeast expression vector without generating intermediate connector vectors, we designed the **pCas9-*GFP*^{MX}** (CEN6-*URA3*) vector. Primers were designed to PCR amplify the *GFP* expression cassette (from pYTK047) with GG enzyme sites to clone *GFP* (BsaI). The BsmBI sites, with overhangs similar to ConLS and ConRE, designed in the primers, allow the PCR-based sgRNA TU cloning in CEN6-*URA3* yeast expression vector (Table S1). The primers used to amplify the sgRNA expression units harbor connector overhangs to clone >1 sgRNA expression unit in tandem using GG protocol, eliminating the need to individually make each connector clone.

Repair template design and construction

Repair DNA was designed to be a linear DNA molecule that contained the human or carotenoid gene coding sequence, from the start codon to stop, with at least 80-100bp of flanking homology to the yeast genome immediately upstream and downstream of the native yeast start and stop codons or landing pad loci (*511B*, *USERX*, *FGF20*).[49] The repair templates were constructed by PCR using ORFeome[50] or MGC[51] clones or carotenoid gene plasmids,[47] using primers with long extensions providing the homology sequence. Repair template PCRs were performed with Accuprime Pfx (Thermo) as multiple 100μL reactions according to the manufacturer's protocol, combined and purified using the Zymo DNA Clean or Qiagen PCR Cleanup and Concentrator −25 kit.

Clone verification

Clones were initially screened by colony PCR using a rapid DNA isolation method and colony PCR.[52] Forward primers for PCR screening were designed such that the upstream primer would bind in the yeast genome approximately 150bp - 500bp upstream of the yeast ORF or the landing pad loci to ensure it was outside of the homology region used for repair. Two reverse primers were designed for each locus, one binding into the yeast and the other within the human or carotenoid gene sequence. To facilitate multiplex PCR screening, each primer pair amplifies different-sized bands for the yeast and human or carotenoid genes. Following plasmid loss, clones were further verified by Sanger sequencing.

Plasmid loss assays

Successful clones were subjected to a plasmid loss procedure to alleviate any stress incurred by the constitutive expression of the Cas9 protein and allow successive transformation of the knockout plasmids. Clones were grown overnight in YPD and then spread again on YPD (100μL of 1:1000 dilution) and replica plated on SC -Ura and YPD or patch plated (typically 6–12 colonies) to both minimal media lacking uracil and YPD. This procedure resulted in around 10–60% of colonies losing the plasmid (estimated). Strains that eliminated URA3-plasmids were selected using 5-FOA.

Growth assays

Cells were diluted to approximately 0.01–0.02 OD660 ($\sim$2-5 x10^5 cells) in 150μL YPD across 3–4 replicates. They were grown and read in a BioTek Synergy H1 or Sunrise Tecan microplate reader, with continuous double orbital shaking at 30°C or 37°C. Reads were taken while stationary every 10–20 min, and experiments were run for at least 24–36 h. For spot assays, strains were grown overnight in YPD at 30°C and spotted with serial dilutions on YPD agar. The plates were incubated for 2–3 days at 23°C, 30°C and 37°C.

CRISPR plasmid based selection to select unique genotypes

Haploid *MATα* engineered yeast strains were each cultured overnight in YPD separately and back-diluted into a pool the next day at equal OD and grown to the mid-log phase. Competent cells were prepared and transformed with various **pCas9-sgRNA^{locus}** plasmids. We performed the experiments as biological and technical replicates using different genotypes in the mix. Competent cells were generated using Zymo Research Frozen EZ kit transformation protocol. Each transformation was plated on SD−URA. Colonies from each Petri plate were picked randomly and genotyped by locus-specific PCR to confirm that CRISPR plasmid exclusively selected a unique genotype.

MERGE^{Q} assays

The engineered strains with single gene modification in BY4741 (Haploid *MATα*) backgrounds were mated with SGA-haploid strains harboring haploid-specific markers (SGA *MATα*; *can1Δ::STE2pr_Sp_HIS3,his3Δ1leu2Δ0ura3Δ0;* select on SD−His+Canavanine & SGA *MATα*; *lyp1Δ::STE3pr_LEU2 his3Δ1leu2Δ0ura3Δ0;* select on SD−Leu+Thialysine). Each parental strain was transformed with two different empty vector plasmids with distinct selectable markers *URA3* and *kanMX*, grown overnight in 5mL media (SC -Ura or YPD +G418). The haploid strains were mixed in a rich medium and incubated at 30°C for 3–4 h with shaking. 500μL of the mated mixture was washed with distilled water, and 10μL of the mix was plated at different dilutions on solid agar media using SD−URA+G418 selection and incubated for 2–3 days at 30°C. Plasmids were cured and the heterozygous colonies were confirmed

by PCR genotyping. The confirmed heterozygous strains for all humanized proteasome and carotenoid strains were transformed with CRISPR plasmids targeting wild-type loci. Several colonies were picked randomly from the plate and genotyped for homozygosity using the locus specific PCR.

Alternatively, instead of an empty vector, the engineered humanized or carotenoid strains were transformed with a CRISPR plasmid for which the corresponding strain harbors a CRISPR-resistant locus. The subsequent mating with an opposite mating-type strain with an empty vector (pCas9 alone or a backbone plasmid) results in the loss of heterozygosity and conversion of single loci. In each scenario, the number of colonies obtained on plates transformed with CRISPR plasmid ($\mathbf{CFU_O}$) were compared with empty vector transformations ($\mathbf{CFU_E}$) to calculate the efficiency of **MERGE^0**. Each experiment was performed at least 3 times.

Estimating the efficiency of MERGE^0 using yeast ADE2 locus readout

The **pCas9-sgRNA^{ADE2}** was transformed in a diploid hetKO *ADE2/ade2Δ::kanMX* obtained from the yeast "Magic Marker" hetKO collection.[28] The CRISPR plasmid transformation resulted in no lethality with all red colonies. To verify that the red colonies observed after **MERGE^0** are not due to the mutations in the *ADE2* locus (NHEJ) and instead due to conversion to *ade2Δ::kanMX* locus (HDR), the heterozygous diploid (ADE2/*ade2Δ::kanMX*) strain transformed with a control (**pCas9 alone**) and **pCas9-sgRNA^{ADE2}** plasmids were sporulated followed by tetrad dissection. Haploid spores were selected on YPD or YPD + G418 (200 μg/mL).

Compared to **MERGE^0**, the co-transformation of **pCas9-sgRNA^{ADE2}** and oligo as a repair template harboring (100bp homology, 5X molar excess than the plasmid) to the 5′ AND-3′ UTRs of *ADE2* locus in haploid wild-type yeast cells resulted in significantly fewer survivors (% $\mathbf{CFU_O/CFU_E}$ = 21.6± standard deviation; N = 4). However, this method is still far less-efficient than **MERGE^0** (∼100% vs. 21.6%).

Sporulation of diploid and selection of haploid yeast strains

Diploid strains were sporulated in a media containing 0.1% potassium acetate and 0.005% zinc acetate for 4–7 days. For Tetrad dissection, spores were spun down at 5000rpm for 5 min, resuspended in 200μL of 20 mg/mL Zymolyase and incubated at 37°C for 25–30 min. The mix was then incubated at −20°C to stop the reaction. Cells were thawed on ice, and 20μL of the mix was plated into YPD for Tetrad dissection. Tetrad Dissection was performed using Spore play (Singer Instruments) on YPD and then replicated on SGA selection. For carotenoid and humanized loci in SGA background strain, the sporulation mix was directly plated on SGA selection. Locus-specific PCR verified the haploid strains for the presence or absence of engineered loci.

Using MERGE^0 to perform one-step gene essentiality assays in yeast

HetKO diploid strains for 7 α-proteasome core genes were obtained from the yeast "Magic Marker" hetKO collection.[28] The strains were transformed with either a single-sgRNA-CRISPR plasmid targeting the corresponding yeast α-proteasome genes or the empty vector control and selected on SD−URA+G418. CRISPR plasmid transformed plates caused lethality in 6 of 7 α-proteasome hetKO strains, except in the case of non-essential *α3*. A similar assay in the hetKO *ADE2/ade2Δ::kanMX* strain showed viable homozygous null cells for a non-essential *ADE2* locus.

MERGE^0 assays across yeast chromosome I

All the strains harboring heterozygous knockout diploid loci from the yeast "Magic Marker" collection on chromosome I were arrayed in a 96-well format.[28] The sgRNA targeting the *kanMX* cassette was designed using the strategy mentioned earlier (Table S1). Since the bacterial selection for the yeast shuttle vector harbors a *Kan^R* cassette identical to the *kanMX* cassette in hetKO strains, we designed an expression vector with *AmpR* selection.

Each strain was inoculated in 800μL of YPD + His(50 mg/L) + G418 (200 μg/mL) overnight. The following day, cultures were back-diluted and grown to the mid-log phase. Competent cells were generated in a 96-well format using the Gietz yeast transformation protocol.[53] Each strain was transformed with either 1μg of the pCas9 alone or **pCas9-sgRNA^{kanMX}** plasmid. An equal amount of the transformed cells were spotted on SD−URA and SD−URA+G418 medium.

MERGE$^{M\&M}$ assays to combine 2 loci using the mate & mix strategy.

Each haploid strain (of mating types) 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 h 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 the 1 μg of pCas9 alone or double-sgRNA-CRISPR plasmid followed by selection and PCR genotyping.

MERGE^1 and MERGE^{MX} assays at 2 or more loci using humanized proteasome and carotenoid yeast strains

For ***MERGE^1*** or ***MERGE^{MX}***, each double- or multiple-sgRNA-CRISPR plasmid was tested to show lethal phenotype the wild-type or single-humanized strains. Transformation of 0.5-1μg of double- or multiple-sgRNA-CRISPR plasmid in heterozygous diploid strains either showed no lethality or a lethal phenotype (depending on whether genotype was compatible). Genotypes were confirmed by performing a locus-specific PCR.

CRISPR plasmid *assay to target multiple GFP loci*

To address the scalability of targeting multiple yeast loci, we used the Green Monster (GM) strain.[19] **pCas9-sgRNA^{GFP}** [*GFP* (S65T)] was constructed using the direct cloning strategy in a backbone with *Amp^R* (*E. coli*) and *kanMX* (yeast) selection. The *MAT*α mating-type Green Monster was used [*MAT*α*; lyp1*Δ *his3*Δ*1 leu2*Δ*0 ura3*Δ*0 met15*Δ*0 can1*Δ*::GMToolkit-*α *[CMVpr-rtTA NatMX4 STE3pr-LEU2]* for the assay. The OFF-target activity of **pCas9-sgRNA^{GFP}** was tested in a haploid WT strain with no *GFP* gene showing a **CFUo/CFU$_E$ =**$\sim$ **1**. The ON-target activity of the CRISPR plasmid was measured by transformation in natively tagged C-terminal fusion of *GFP* to *BRO1* gene *(Bro1-GFP)* obtained from the yeast *GFP* collection.[54]

Cells were grown in YPD overnight and back diluted in SC media with 10 μg/mL of doxycycline for $\sim$48 h. The culture was diluted 1 in 10 in water and passed through the flow cytometer (BD Accuri C6 Plus). In the case of GM cells transformed with **pCas9-sgRNA^{GFP}**, several surviving colonies were picked, pooled and the mixture passed through the flow cytometer.

MERGE^{MX} assays to assemble an entire carotenoid pathway in yeast

SGA strains of both mating-types harboring single-carotenoid transcription units at the landing pad loci were cultured independently overnight, followed by mixing at 0.3 OD. A hole was made on the lid of the Eppendorf tube to allow aeration, and the mixture was incubated at 30°C for 4–6 h. The formation of diploids was confirmed by light microscopy. Next, 500μL of culture was centrifuged at 3500 rpm for 5 min. The supernatant was removed and resuspended in 2 - 5mL of sporulation media, followed by incubation on a rotating shaker for 5–8 days at room temperature. Next, 500 μL of sporulation mix was centrifuged at 3500 rpm for 5 min and treated with Zymolyase. After the first sexual cycle, the cells were centrifuged, washed with water, and resuspended in YPD. The cycle was repeated to generate diploids. The mixture was incubated for 4–6 h in a shaking incubator at 30°C for 2 days. The confirmation and appearance of most diploids were confirmed with a light microscope, and the 100μL of culture was inoculated overnight to make competent cells. The competent cells were transformed with either an empty vector or a triple-sgRNA-CRISPR plasmid **pCas9-sgRNA^{511B,USERX,FGF20}**. For the most part, we followed a *GFP* monster protocol. However, we did not use haploid or diploid-specific selection.[55] Instead, CRISPR selection was sufficient to enrich unique combined genotypes.

Whole genome sequencing (WGS)

Genomic DNA was purified using the Monarch Nucleic acid purification kit according to the manufacturer's protocol (NEB). Spheroplasts were obtained before the genomic DNA extraction for a high-quality DNA prep. Wild type and engineered strains were sequenced using Illumina MiSeq 2x150 at 30x coverage using 150-bp paired-end reads. The Geneious Pro Software[48] and its included tools were utilized for pairing paired-end sequences, trimming ends and adapters based on quality using the BBDuk tool, and reference mapping using the Geneious Read Mapper algorithm at medium-low sensitivity. Reads were mapped to a reference BY4741 strain and its modified version by replacing the sequences of the 5 engineered genes with their human ortholog to ensure alignment with the humanized loci reads. Mapping was run at medium sensitivity: word length was set to 18 with a maximum permitted mismatch of 20% of the read length, maximum mismatches per read were set to 20%, a minimum 80% overlap was required, and reads with errors were set to accurately be mapped to repeat regions - this was iterated 5 times to give the final mapping. For analysis, low coverage regions of below 2 standard deviations from the mean were excluded from SNP-calling, and only SNPs with a variant frequency of 0.90 or higher were considered. SNPs were called with a minimum variant frequency of 0.25, and low coverage regions below 2 standard deviations from the mean were excluded. SNPs that were unique to the engineered strain (i.e. not in the mappings of the wild-type strain) were marked. To ensure SNPs were not introduced by off-target CRISPR effects, 300bp up- and downstream of each SNP was searched for CRISPR-gRNA alignment at 75% sequence similarity.

Construction of yeast expression vectors harboring multiple human β *proteasome genes*

Human β subunits, with flanking Type IIS restriction sites generating unique overhangs, were obtained as gBlocks from IDT. To construct a yeast expression vectors for native expression of 7 human constitutive- (pCN7) and immuno-βs (pIN7), we PCR amplified 5′UTRs ($\sim$500bp) and 3′UTRs ('250bp) of the orthologous yeast genes using genomic DNA (BY4741) as a template. Each 5′UTR, gBlock (human β ortholog) and 3′UTR were cloned as transcription units (TUs) using Type IIS (Golden Gate)-based cloning. Each TU was sequence-verified followed by cloning 7 constitutive- and immuno-β TUs in yeast vector (CEN6, URA+).

To obtain a yeast expression vectors for heterologous expression of 5 functionally non-replaceable human constitutive- (pCH5) and immuno-βs (pIH5),[26,45] the human β gBlocks were cloned as transcription units (TUs) using strong yeast promoters and terminators from MoClo toolkit and Type IIS (Golden Gate)-based cloning.[17] The TUs were sequence-verified followed by cloning 5 constitutive- and immuno-β TUs in yeast vector (CEN6, URA+).

Proteomic sample preparation

The cell pellet obtained from the strains was resuspended in Digestion Buffer (50mM Tris, 2mM $CaCl_2$) and lysed by bead beating with glass beads for 3 x 1 min. The resulting whole-cell lysate was recovered and mixed 1:1 with 2,2,2-trifluoroethanol. Samples were reduced by incubation with 5 mM tris(2-carboxyethyl)phosphine (TCEP solution, Pierce) at 60°C for 40 min. Reduced samples were then alkylated by incubation with 15 mM iodoacetamide at room temp. for 30 min. 7.5mM dithiothrietol (DTT) was added to quench excess iodoacetamide. Samples were then diluted 10-fold with Digestion Buffer and 2 μg trypsin was added for digestion. Samples were incubated at 37°C for 5 h. Tryptic digestion was quenched with 1% formic acid, and samples were concentrated by

vacuum centrifugation to reduce volume to less than 300 μL. Digests were cleaned using Hyper-Sep C18 SpinTips (Thermo) according to manufacturer's protocol. Eluted peptides were briefly dried by vacuum centrifugation, then resuspended in 5% acetonitrile, 0.1% formic acid.

LC-MS/MS analysis

Tryptic peptides were separated by reverse phase chromatography on a Dionex Ultimate 3000 RSLCnano UHPLC system (Thermo Scientific) with an Acclaim C18 PepMap RSLC column using a 3–42% acetonitrile gradient over 60 min. Peptides eluted directly into a Thermo Orbitrap Lumos mass spectrometer by nano-electrospray. Data-dependent acquisition was applied, with precursor ion scans (MS1) collected by FTMS at 120,000 resolution and HCD fragmentation scans (MS2) collected in parallel by ITMS with 3-s cycle times. Monoisotopic precursor selection and charge-state screening were enabled, with ions > +1 charge selected. Dynamic exclusion was applied to selected ions +/− 10 ppm for 30 s.

Raw data was processed using Proteome Discoverer 2.2 (Thermo Scientific). Mass spectra were searched against a protein sequence database comprising the *Saccharomyces cerevisiae* reference proteome (UniProt OX: 559292), human proteasome proteins, and a list of common protein contaminants (MaxQuant). Searches were restricted to fully tryptic peptides only, allowing up to two missed cleavages. A precursor tolerance of 5 ppm and fragment mass tolerance of 0.5 Da were used. Static modifications of carbamidomethyl cysteine and dynamic modifications of oxidized methionine and protein N-terminal acetylation and/or methionine-loss were considered. High-confidence peptide-spectrum matches (PSMs) were filtered at a false discovery rate of <1% as calculated by Percolator. Peptide abundances based on extracted ion-chromatography (XIC) feature intensities were calculated using the Label-Free Quantitation (LFQ) workflow.

Due to conservation of sequence between human and yeast proteasome proteins, peptide and protein abundances were calculated using an in-house workflow to remove degenerate peptides. Isoleucine/leucine sequence variants were collapsed into single peptide groups. For each peptide, protein groups generated by Proteome Discoverer 2.2 were checked to make sure all peptides used were unique to either yeast or human protein. Protein abundances were calculated as the sum of all unique peptides matching to that protein. Only proteins detected in two separate injections of at least one sample were included in the final dataset. The consolidated data is provided as Table S2.

Microscopy

The yeast cells were grown to mid-log phase in selective media to maintain plasmids and were imaged with DMi6000B microscope (Leica Microsystems). Cell size measurements were performed with a minimum 50–100 cells per strain using FIJI/ImageJ.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data were analyzed using GraphPad Prism version 9.0.1, GraphPad Software, San Diego, California USA, and Excel Office (Microsoft, WA, All the statistical details of experiments can be found in the figure legends and results, including the statistical tests and number of replicates for each experiment. Figures were made with Adobe Illustrator 2020 (Adobe Systems, CA, USA). Mass spectrometry data analysis was performed using Proteome Discoverer 2.2 (ThermoFisher) and heatmaps were generated using MORPHEUS software (https://software.broadinstitute.org/morpheus/).