Development of a multi-locus CRISPR gene drive system

in budding yeast

Yao Yan[1] and Gregory C. Finnigan[1*]

[1]Department of Biochemistry and Molecular Biophysics, Kansas State University, 141 Chalmers

Hall, Manhattan, KS 66506 USA

*Correspondence to: Gregory C. Finnigan, Phone: (785) 532-6939; FAX; (785) 532-7278;

E-mail: gfinnigan@ksu.edu

ABBREVIATIONS

CRISPR, clustered regularly interspaced short palindromic repeats; NHEJ, non-homologous end

joining; HR, homologous recombination; DSB, double stranded break; sgRNA, single guide RNA

fragment; GD, CRISPR-based gene drive system; CGM, "complete" multi-locus drive; MGD,

"minimal" multi-locus drive; BRCT, BRCA1 C-Terminal domains; UTR, untranslated region;

indel, insertion or deletion.

ABSTRACT

The discovery of CRISPR/Cas gene editing has allowed for major advances in many biomedical disciplines and basic research. One arrangement of this biotechnology, a nuclease-based gene drive, can rapidly deliver a genetic element through a given population and studies in fungi and metazoans have demonstrated the success of such a system. This methodology has the potential to control biological populations and contribute to eradication of insect-borne diseases, agricultural pests, and invasive species. However, there remain challenges in the design, optimization, and implementation of gene drives including concerns regarding biosafety, containment, and control/inhibition. Given the numerous gene drive arrangements possible, there is a growing need for more advanced designs. In this study, we use budding yeast to develop an artificial multi-locus gene drive system. Our minimal setup requires only a single copy of *S. pyogenes* Cas9 and three guide RNAs to propagate three separate gene drives. We demonstrate how this system could be used for targeted allele replacement of native genes and to suppress NHEJ repair systems by modifying DNA Ligase IV. A multi-locus gene drive configuration provides an expanded suite of options for complex attributes including pathway redundancy, combatting evolved resistance, and safeguards for control, inhibition, or reversal of drive action.

INTRODUCTION

The discovery and implementation of CRISPR/Cas gene editing has revolutionized countless fields and sub-specialties across molecular biology and biotechnology to improve human health, agriculture, ecological control, and beyond. Briefly, alteration of the genetic code is accomplished using (i) a bacterial derived nuclease (typically Cas9 or Cas12a), (ii) a single-stranded fragment of "guide" RNA, and (iii) an optional exogenous repair fragment of DNA (JINEK *et al.* 2012; JINEK *et al.* 2013; DOUDNA AND CHARPENTIER 2014; ZETSCHE *et al.* 2015). Priming of the nuclease with a pre-programmed guide RNA fragment targets a specific genomic sequence for a double stranded break (DSB). Following DNA cleavage, eukaryotic cells activate repair systems to either fuse broken chromosomal ends together via non-homologous end joining (NHEJ) or, in the presence of donor DNA, introduce exogenous sequence via homologous recombination (HR). Moreover, the CRISPR methodology is not restricted to DSB-induced alteration of the genome—recent efforts have demonstrated that nuclease-dead variants (e.g. dCas9) can serve as delivery systems to modulate transcriptional activity (QI *et al.* 2013), alter epigenetic landscapes (THAKORE *et al.* 2015), or introduce mutational substitutions *sans* any DNA cleavage event (GAUDELLI *et al.* 2017).

One powerful biotechnological application of the CRISPR methodology is within a "gene drive" system. The basic design of a homing drive includes the expression constructs for the CRISPR nuclease and the corresponding guide RNA positioned at a desired locus of choice—the mechanism of propagation involves targeting of the homologous chromosome (within a diploid or polyploid organism) at the same genetic position (typically cleaving the wild-type gene). Creation of a DSB followed by HR-based repair (using the gene drive-containing DNA as a donor) causes the entire artificial construct (Cas9, the sgRNA, and any desired "cargo") to be copied; in this way,

3

a heterozygous cell is automatically converted to the homozygous state. This *super*-Mendelian genetic arrangement allows for the forced propagation of a genetic element within a population and has the potential to modify entire species on a global scale (BULL AND BARRICK 2017; GODFRAY *et al.* 2017). Some of the possible benefits of this technology include eradication of invasive species (ESVELT AND GEMMELL 2017; PROWSE *et al.* 2017), agricultural pest management (COURTIER-ORGOGOZO *et al.* 2017), and elimination of insect-borne diseases such as malaria (GODFRAY *et al.* 2017; HAMMOND AND GALIZI 2017; LAMBERT *et al.* 2018). A number of recent studies have demonstrated the potency and success of CRISPR-based gene drives in fungi (DICARLO *et al.* 2015; ROGGENKAMP *et al.* 2017; ROGGENKAMP *et al.* 2018; SHAPIRO *et al.* 2018), and metazoans (GANTZ *et al.* 2015; HAMMOND *et al.* 2016; CHAMPER *et al.* 2017; GRUNWALD *et al.* 2018). While ongoing technical challenges remain in the design, optimization, and field testing of gene drive-harboring organisms, there are also serious biosafety and ethical concerns regarding use of this biotechnology as even current drive systems are expected to be highly invasive within native populations (NOBLE *et al.* 2018). There is an immediate need for further study (*in silico* and *in vivo*) of gene drive systems that focus on issues of safety (DICARLO *et al.* 2015; NAJJAR *et al.* 2017; JAMES *et al.* 2018), control and reversal (VELLA *et al.* 2017; BASGALL *et al.* 2018), and optimal design (PROWSE *et al.* 2017).

There are many types of gene drive designs including "daisy-chain drives," "underdominance drives," and "anti-drives," each with a distinct arrangement of the basic CRISPR components that is predicted to sweep through native populations at varying levels/rates (GODFRAY *et al.* 2017; BURT AND CRISANTI 2018; DHOLE *et al.* 2018). Moreover, the need for additional drive components (more than one guide RNA construct), genetic safeguards, and built-in redundancy, calls for a new level of complexity within drive architecture. Here, we demonstrate

4

use of *multiple gene drives* across three chromosomal loci within an artificial budding yeast system. Our "minimal" multi-locus gene drive arrangement requires only a single copy of the *S. pyogenes* Cas9 gene (installed at one position), along with three distinct guide RNAs to multiplex the nuclease throughout the genome. We demonstrate that this technique could be used to perform targeted replacement of a native gene (under its endogenous promoter) *in trans* from the Cas9-harboring locus. Finally, reducing or modulating NHEJ by targeting the highly conserved DNA Ligase IV may provide a means to further bias HR-dependent repair and action of gene drives across eukaryotic systems. Our method includes multiple layers of genetic safeguards as well as recommendations for future designs of multi-locus drive systems.

RESULTS

Rationale and design of a multi-locus CRISPR gene drive

To date, a number of studies in fungi, insects, and now vertebrates, have demonstrated that CRISPR-based gene drive systems are effective in both single-celled and multicellular eukaryotes (DICARLO *et al.* 2015; GANTZ *et al.* 2015; HAMMOND *et al.* 2016; CHAMPER *et al.* 2017; GRUNWALD *et al.* 2018; ROGGENKAMP *et al.* 2018; SHAPIRO *et al.* 2018). One of the benefits of homing systems is the ability to install additional genetic "cargo" proximal to the gene drive (consisting of a nuclease gene and an expression cassette for the guide RNA). Current strategies use the gene drive cassette itself to delete and replace an endogenous gene, and/or include exogenous material as a desired cargo. However, there are a number of limitations to the use of a single locus harboring the entirety of the gene drive. First, addition of entire genetic pathways or large numbers of gene expression systems may be less efficient at HR-based copying of the drive. Second, introduction of additional endogenous gene(s) or modified alleles may require the native promoter system and/or epigenetic landscape to provide accurate and timely expression—this

5

would not be possible at a single generic drive-containing locus. Third, given the observation of both natural (e.g. single nucleotide polymorphisms) and evolved resistance to gene drives (indels resulting from NHEJ) within insect populations (CHAMPER *et al.* 2017; DRURY *et al.* 2017; HAMMOND *et al.* 2017; UNCKLESS *et al.* 2017; BUCHMAN *et al.* 2018; KARAMINEJADRANJBAR *et al.* 2018), mechanisms for fortifying drive systems are still being elucidated. The proposal to increase the number of targeted double stranded breaks (and corresponding sgRNAs) to the single nuclease of choice (e.g. *S. pyogenes* Cas9) would greatly aid in combatting resistance (BULL AND MALIK 2017; MARSHALL *et al.* 2017; NOBLE *et al.* 2017). However, an independent means to both minimize or escape resistance *and* ensure the intended biological outcome (deletion of the intended gene or introduction of the exogenous cargo) would involve a redundant delivery system. In this way, multiple gene drives (with multiple guide RNAs) within the same organism could target independent genetic loci either from the same, distinct, or parallel genetic pathways to achieve the desired outcome(s).

We envisioned two general strategies for the development of a gene drive system across distinct chromosomal positions: (i) each multi-locus "Complete" Gene Drive (CGD) would contain both a nuclease and corresponding guide RNA or (ii) a multi-locus "Minimal" Gene Drive (MGD) would include a nuclease and sgRNA, and all other genetic loci would *only* contain additional guide RNA cassettes (Fig. 1A, *left*). We chose to focus on the latter strategy for a number of reasons, but we recognize that both would have distinct challenges and advantages. For one, a possible technical hurdle to development of a modified organism with multiple CGDs would be the generation of distinct "large" expression system consisting of the entire nuclease gene, flanking UTR, the guide expression cassette(s), and any optional cargo compared to the MGD which removes the bulk of the drive system (nuclease expression) at additional loci.
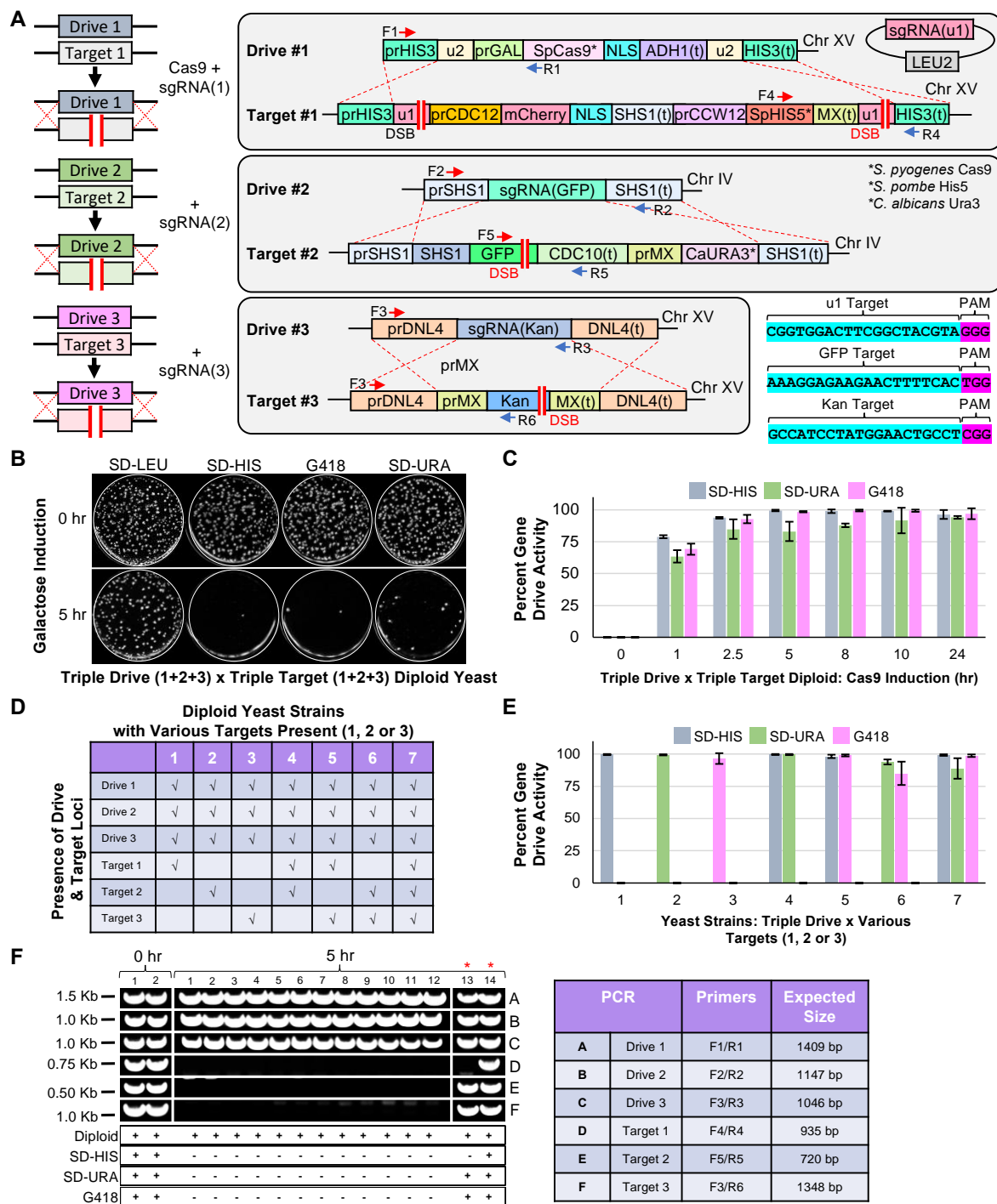
Figure 1. Design of a CRISPR/Cas9-based gene drive system in *S. cerevisiae* across three loci.

(A) *Left,* An artificial gene drive was installed at three separate loci in haploid yeast. Each drive

system (Drive 1-3) contained a guide RNA cassette targeting an artificial target (Target 1-3) at the

7

same locus. Only Drive 1 contained the cassette for *S. pyogenes* Cas9. *Right,* Artificial (u1) and (u2) sites (FINNIGAN AND THORNER 2016) were used flanking the gene drive at the *HIS3* locus (Chromosome XV) and the *S.p.HIS5* selectable marker. The *SHS1* locus (Chromosome IV) included a C-terminal GFP fusion and the *C.a.URA3* marker. *DNL4* (Chromosome XV) was deleted with the Kan^R cassette. All sgRNAs were targeted to non-native sites (sequences shown). The sgRNA(u1) cassette was included on a high-copy plasmid (*LEU2* marker). *S.p.*Cas9 was under control of the inducible *GAL1/10* promoter. (B) Haploid yeast harboring the triple drive (GFY-3675) were mated to yeast of the opposite mating type containing the triple target (GFY-3596) to form diploids. Cas9 expression was induced by growth in galactose for either 0 or 5 hr. Yeast were diluted to 100-500 cells per plate, grown for 2 days, and transferred to SD-LEU, SD-HIS, SD-URA, and G418 plates. (C) A time course of galactose activation using the [GFY-3675 x GFY-3596] diploid in triplicate. Error, SD. (D) Seven haploid yeast strains (GFY-3206, 3593, 3264b, 3578, 3594, 3623, and 3596) were constructed and tested as in (B) against the triple drive strain (GFY-3675). (E) Each of the diploids created from (D) were cultured for 5 hr and quantified for drive success. Error, SD. (F) Clonal isolates were obtained from diploids generated in (B) at either 0 hr (2 isolates) or 5 hr activation of Cas9 (14 isolates). All yeast were confirmed as diploids and assayed on each media type (*below*). Diagnostic PCRs were performed on genomic DNA to detect the presence (or absence) of each drive and target locus; oligonucleotide (Supplementary Table 5) positions can be found in (A) and the expected PCR fragment sizes are illustrated (*right*). Two isolates (13,14) were chosen for their incomplete growth profile (red asterisks).

Along these lines, the issue of appropriate expression of each of the nuclease gene(s) (whether identical or distinct) would need to be addressed using identical or modified promoter

elements; this issue does not exist for a MGD with only a single copy of Cas9. Second, the issue of biosecurity and safeguarding against accidental or malicious release was taken into consideration. Given that a MGD would only harbor one copy of the nuclease, it would provide far less hurdles to counter and inactivate—either through use of an anti-drive system (DICARLO *et al.* 2015; VELLA *et al.* 2017), by induced self-excision (ROGGENKAMP *et al.* 2018), or by removal of the Cas9-containing drive guide RNA (ROGGENKAMP *et al.* 2018). Therefore, we have chosen to focus our study on design and testing of a three-locus MGD in budding yeast using the *S. pyogenes* Cas9 nuclease.

An efficient triple gene drive system functions independently at each locus

Our novel system includes the most potent genetic safeguard known to date used within a gene drive: artificial and non-native sequences used as targets. In this way, we have not only generated a haploid yeast strain harboring the MGD system at three genetic loci (*HIS3, SHS1,* and *DNL4*), but have also created a corresponding haploid strain with three distinct artificial *targets* at the same three loci (Fig. 1A, *right*). The "primary" drive at the *HIS3* locus includes (i) Cas9 under an inducible promoter (*GAL1/10*) commonly used for overexpression, (ii) flanking (u2) artificial sequences to be used for self-excision as a safeguard, and (iii) the absence of any selectable marker. The corresponding guide RNA cassette was installed on a high-copy plasmid for security reasons, but could have also been integrated proximal to the drive itself. The "secondary" and "tertiary" drive systems (*SHS1* and *DNL4*) are both non-essential genes and contain the minimum required components in the MGD design; in both cases, the native gene was deleted and fully replaced by the guide expression cassette (455 bp, although this could be reduced further) with no selectable marker. Construction of this complex haploid yeast strain used a combination of traditional HR-based integrations (with selectable markers), universal Cas9-targeting systems (CRISPR-

9

UnLOCK) (ROGGENKAMP *et al.* 2017), and novel "self-editing" integration events (Supplementary Fig. 1). To test the efficacy of the MGD, a three-locus "target" strain was generated: the *HIS3* locus was flanked by two (u1) sequences and included the *S.p.HIS5* selectable marker, the *SHS1* gene was fused with GFP and contained the *C.a.URA3* marker, and finally, *DNL4* locus was deleted and replaced with the Kan^R drug cassette (Fig. 1A, *right*).

The triple MGD strain was first mated with the triple target strain to form a diploid, and Cas9 was activated by culturing in medium containing galactose (Fig. 1B). In the absence of nuclease expression (Fig. 1B, *top*), all yeast colonies contained the (u1) guide plasmid (*LEU2*), and three selectable markers (*HIS5*, *URA3*, and Kan^R—providing resistance to G418). However, following a 5 hr incubation in galactose, >95% of all colonies were sensitive to all three growth conditions indicating a loss of all three selectable markers and replacement via the MGD (Fig. 1B, *bottom*). A time course of galactose induction illustrated highly efficient drive activity for all three loci by five hours; we noticed a slight lag in efficiency for the loss of the *URA3* marker (*SHS1*) locus until the 24 hr mark (Fig. 1C). This observation may be due to the *HIS3* and *DNL4* loci both being present on chromosome XV whereas *SHS1* was located on chromosome IV. Alternatively, differences in available guide RNAs (plasmid-borne versus integrated) or local epigenetic effects could cause this slight reduction in editing. Next, to ensure that action of the MGD at each locus was not dependent on the presence or absence of one or more of the intended targets (simulating "resistance" at one or more loci), we retested the triple drive strain against six additional strains, each lacking one or two of the proper targets and instead, contained the native yeast sequence: *his3Δ1*, *SHS1*, or *DNL4* (Fig 1D). We obtained similar results for each combination as the triple MGD strain (#7) indicating that each gene drive functioned independent of the presence of additional target(s) (Fig. 1E). We also observed that drive success at the *SHS1* locus slightly

increased when fewer targets were presented. Finally, to ensure that the loss of the selectable marker was coupled to replacement of the target locus by the drive locus, we isolated clonal yeast from the MGD triple cross (Fig. 1B) and confirmed both the growth profile and ploidy status of random samples (Fig. 1F, *bottom*). Diagnostic PCRs were performed on all six distinct loci to assay for the presence or absence of each engineered drive and target (Fig. 1F, Supplementary Fig. 3). Oligonucleotides (Supplementary Table 5) unique to specific drive/target elements were chosen; prior to Cas9 activation (0 hr), diploids contain all six distinct loci (two isolates). However, following activation of the nuclease, diploids maintained all three drive loci (PCRs A,B, and C), but lost all three target loci (PCRs D, E, and F) (twelve independent isolates).

We recognized that following the 5 hr drive activation, a small number of yeast colonies (<5% in most cases) still contained one or more selectable marker(s). We reasoned that these rare colonies likely arose from either complete or partial failure of the gene drive system for various possible reasons (poor expression, loss of guide RNA plasmid, NHEJ, alterations in ploidy, etc.). Therefore, we isolated and tested additional clones that displayed incomplete growth profiles across the three selection plates (isolates 13,14) (Fig. 1F). One isolate (13) had lost the (u1)-flanked target at the *HIS3* locus yet still contained the *SHS1* and *DNL4* markers. The second isolate (14) appeared to have lost the *LEU2*-based plasmid and all three target loci were still present (Fig. 1F). Following transformation with the (u1) guide vector, we examined a second round of drive activation from these two isolates and obtained a similar growth profile with a loss of the remaining loci indicating that at least some of the "failed" drive occurrences resulted from improper activation and/or targeting (Supplementary Fig. 4). Of note, our gene drive system was activated in the absence of any selection—diploids were grown in rich medium containing galactose, and

11

grown on SD-LEU plates prior to testing of the drive status on various medium. In this way, the

action of the gene drive was performed in the absence of any selection or challenge.

Figure 2. DNA Ligase IV, critical for NHEJ and conserved across eukaryotes, provides a unique candidate for gene drives. (A) Phylogenic analysis of Ligase IV candidates (Supplementary Table 3) across fungi and metazoans by Phylogeny.fr (DEREEPER *et al.* 2008; DEREEPER *et al.* 2010). Branch lengths correspond to the number of substitutions per site and the confidence of most branches is illustrated as a decimal (red text). (B) *Top,* Illustration of the domain structure of yeast Dnl4. The catalytic N-terminal portion includes a DNA binding domain, adenylation domain, and oligonucleotide domain (blue). The C-terminal portion includes tandem BRCA1 C-Terminal domains (BRCT). *Bottom,* A multiple sequence alignment was performed

using Clustal Omega (LI *et al.* 2015) of the yeast, mosquito, and human Ligase IV protein C-terminii. Identical residues are shown against a black background and similar residues are colored in blue. Secondary structures (pink cylinder, α-helix; green arrow, β-strand) for the yeast Dnl4 C-terminal as determined by the crystal structure are illustrated (DORE *et al.* 2006). The position of six alleles (K742, T744, L750, D800, G868, and G869) are also illustrated (red asterisk) that were identified from a previous study (CHIRUVELLA *et al.* 2014). (C) The protein sequences of the *A. gambiae* (645-914) and *H. sapiens* (656-911) Ligase IV were modeled against the crystal structure of the *S. cerevisiae* (683-939) Dnl4 (PDB:1Z56) using I-TASSER (ROY *et al.* 2010) (Supplementary Table 4) and illustrated using Chimera (PETTERSEN *et al.* 2004). (D) Cas9-based genomic integration methodology for introduction of mutational substitutions to the native *DNL4* locus in yeast. Two sgRNA-expressing cassettes were cloned onto high copy plasmids (marked with *LEU2* and *URA3*) to induce two DSBs within the C-terminus of *DNL4*. Silent substitutions were generated within the intended repair DNA to prevent re-targeting of Cas9 (silent alterations in yellow). Two repair strategies were used to include either a non-native terminator coupled with a sgRNA(Kan) cassette, or the native *DNL4* terminator; the included amount of homology (bp) is illustrated.

DNA Ligase IV as a target for gene drives

Our choice of the yeast *DNL4* gene as one of the MGD targets was intended to highlight the ability of a drive itself to modify or eliminate non-homologous end joining (NHEJ)—the DNA repair process that directly counteracts the action of gene drives. Following DSB formation by Cas9, the function of the homing drive requires repair of the broken chromosome via homology directed repair using the homologous chromosome (and drive itself) as the source of the donor

13

DNA. However, should NHEJ repair systems ligate the broken chromosome ends prior to HR-based copying, the drive will fail to copy; in fact, imprecise repair by NHEJ may even generate alleles of the target that would be resistant to further rounds of editing. Therefore, this competing DNA repair system remains one major technical hurdle to optimal gene drive design in higher eukaryotes. Of note, interest in modulating, tuning, or inhibiting NHEJ-based repair pathways is not unique to CRISPR gene drives as this mode of repair still competes with the introduction of exogenous DNA via HR (CHU *et al.* 2015; MARUYAMA *et al.* 2015; ROBERT *et al.* 2015; VARTAK AND RAGHAVAN 2015; SCHWARTZ *et al.* 2017; CANNY *et al.* 2018).

The NHEJ pathway is highly conserved from yeast to humans and functions to directly fuse exposed DNA ends (LIEBER 2010; CHIRUVELLA *et al.* 2013b). DNA Ligase IV (Dnl4 in yeast, Lig4 in humans) is required for the final step of DNA ligation along with other conserved binding partners (ELLENBERGER AND TOMKINSON 2008). We examined the genomes of other fungi and metazoans using the yeast Dnl4 protein sequence as a query and a phylogenetic history of this enzyme illustrated the evolution of this enzyme through deep time (Fig. 2A). Note, the branching of *Z. nevadensis* (termite) was poorly supported and has been previously shown to be included within the *Insecta* class (MISOF *et al.* 2014). The DNA Ligase IV enzyme is divided into multiple subdomains including DNA binding, adenylation, oligonucleotide binding, and a C-terminal BRCT domain that interacts with binding partner Lif1 (XRCC4 in human). A previous study identified a number of mutational substitutions within the C-terminus of yeast Dnl4 that resulted in a *partial* loss of function of NHEJ (CHIRUVELLA *et al.* 2014). Examination of protein sequence alignments between yeast, mosquito, and human DNA Ligase IV C-terminal domains revealed only a minor conservation of sequence identity (Fig. 2B). However, several of the identified yeast residues were conserved by either insects and/or humans (yeast T744, D800, G868, and G869).

Using the crystal structure of the C-terminus of yeast Dnl4 as a template, we generated models (I-TASSER) for the corresponding domains of mosquito and human Lig4—both displayed a much higher conservation of structure as opposed to primary sequence (Fig. 2C, Supplementary Table 4). The N-terminal region also displayed strong structural homology using the human Lig4 crystal structure as a template (Supplementary Fig. 5).

While total loss of NHEJ (e.g. *dnl4Δ*) is tolerated in yeast, it is unclear whether a DNA Ligase IV null allele would be viable in higher eukaryotes. Along these lines, truncations or mutations of Lig4 in humans can lead to the rare DNA Ligase IV syndrome (CHISTIAKOV 2010; ALTMANN AND GENNERY 2016). However, given that reduction in transcript or replacement by a partially functioning allele could reduce, but not eliminate NHEJ repair, it could be utilized in other systems to maximize gene drive efficiency, even at the (potential) expense of overall fitness. Therefore, we utilized a "self-editing" methodology to integrate six *dnl4* alleles—five partial loss of function substitutions, one truncation, and a WT control (Fig. 2D). In a strain harboring integrated Cas9 at the *HIS3* locus (ROGGENKAMP *et al.* 2018), we introduced two DSBs within the C-terminus of native *DNL4* and integrated two different constructs: (i) a modified *dnl4* allele with a sgRNA(Kan) cassette and (ii) a modified *dnl4* locus using the native terminator sequence. Both Cas9 target sites were also mutated within the repair (donor) DNA to prevent subsequent rounds of unintended editing.

We utilized these eight haploid strains to quantify the level of NHEJ repair (Fig. 3). Our system of DSB formation followed by DNA repair utilized the dual programmed (u2) sites flanking the Cas9 expression cassette (Fig. 3A). With only a single guide construct, Cas9 would be multiplexed to both identical sites, causing complete excision of the nuclease gene and Kan^R marker. Following transformation of the sgRNA(u2) plasmid, yeast were analyzed for the number

15

of surviving colonies on SD-LEU medium (Fig. 3B). Editing by Cas9 at both (u2) sites followed by precise DNA ligation of the broken ends would generate a "new" (u2) site, and would be subject to a second round of Cas9-dependent cleavage—continual DSB formation followed by exacting repair causes inviability in yeast (ROGGENKAMP *et al.* 2018).
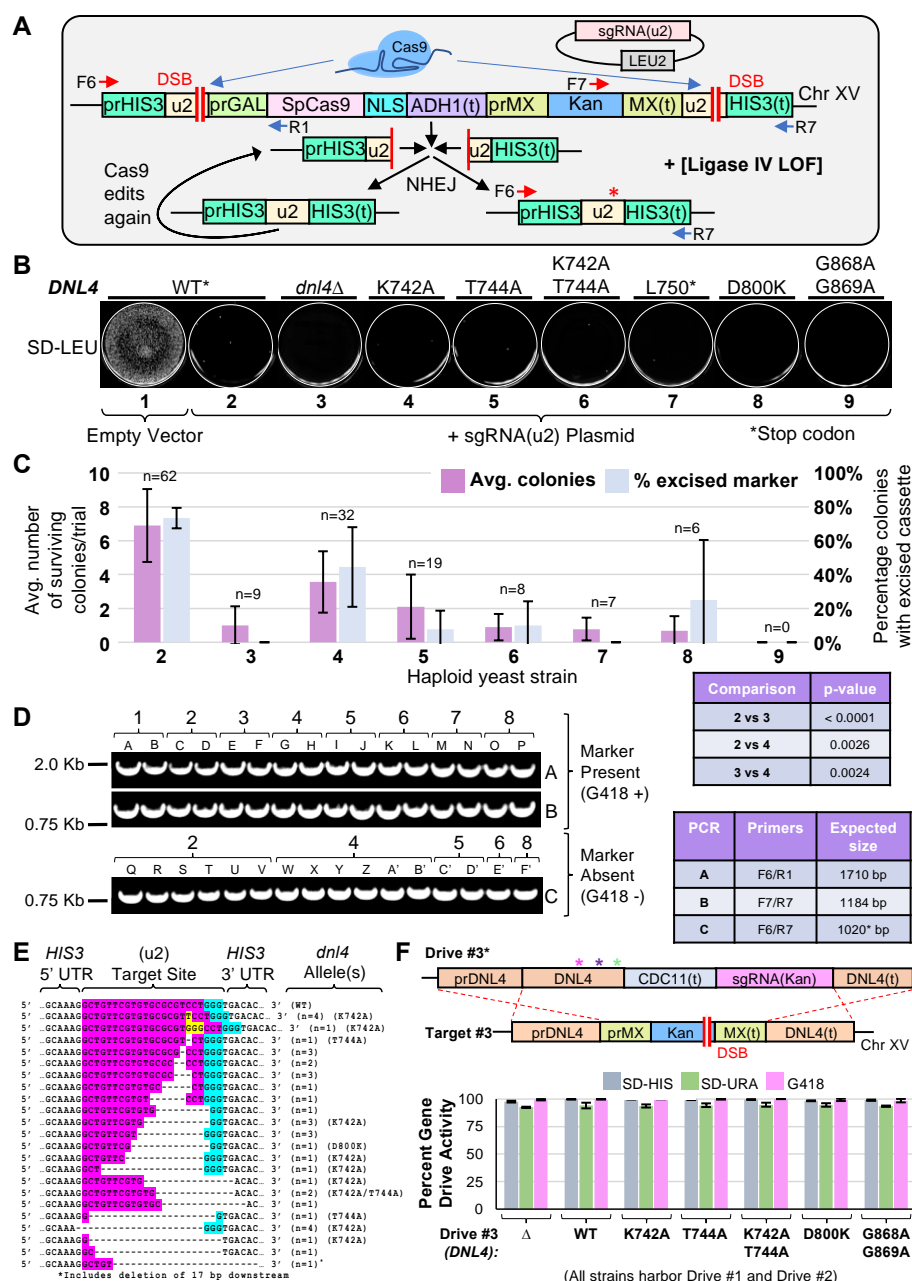

Figure 3. Partial loss of function alleles of yeast DNA Ligase IV reduce NHEJ. (A) Design of a self-excising Cas9-based assay for NHEJ. Strain GFY-2383 included an inducible Cas9 cassette

paired with the Kan^R marker. Transformation of the sgRNA(u2) plasmid would result in multiplexing to two flanking (u2) sites. Repair via NHEJ would result in the formation of the original (u2) site, and would be subject to further rounds of editing; introduction of an indel (red asterisk) would cause destruction of the Cas9 target site. (B) Strains GFY-3850 through GFY-3856 and GFY-3864 (Supplementary Table 1, *Conditions* 1-9) were transformed with the sgRNA(u2) plasmid (pGF-V809) or empty vector control (pRS425) and plated onto SD-LEU for three days. The *DNL4* (WT) gene contained six silent substitutions (asterisk). (C) The average number of surviving colonies was quantified for all trials—labeled as in (B); the number of colonies (n) obtained across all experiments is displayed. Error, SD. For the pRS425 vector, 2897 +/- 357 colonies were obtained. The percentage of isolates that excised the Cas9 cassette at the *HIS3* locus (by sensitivity to G418) is displayed. Error, SD. Statistical analyses of strain comparisons (colonies per trial) were performed using an unpaired t-test. (D) Diagnostic PCRs were performed on chromosomal DNA from clonal isolates from (B) to illustrate presence (2 isolates each) or loss (between 2-6 shown) of the Cas9 cassette. *Conditions* 2-8 correspond to those found in (B). Oligonucleotides (Supplementary Table 5) used can be found in (A) and the expected PCR sizes are illustrated (*right*). (E) DNA sequencing of the *HIS3* locus following NHEJ (on isolates sensitive to G418). For each insertion or deletion, the number of identical clones is displayed. All sequences were obtained from WT *DNL4* yeast except unless otherwise noted. Target, pink. PAM, blue. Insertions, yellow. (F) Triple-drive containing strains were constructed with a modified *DNL4* locus coupled with the sgRNA(Kan) cassette. Haploid drive strains (GFY-3675, 3865-3867, 3871, 3872, and 3875) containing the sgRNA(u1) plasmid were mated with GFY-3596, diploids selected, and gene drives activated as in Fig. 1B. The percentage of colonies sensitive to each condition represented gene drive activity. Error, SD.

However, introduction of an insertion, deletion, or substitution within the target sequence would render the site immune from subsequent rounds of editing. Furthermore, loss of the Kan^R marker provided a growth phenotype associated with targeting of the (u2) sites and excision of the entire cassette at the *HIS3* locus. Both the total number of surviving colonies as well as the percentage of isolates with an excised marker were quantified in triplicate (Fig. 3C). In our assay, the presence of WT *DNL4* allowed for approximately 7 colonies/experimental trial, whereas *dnl4Δ* yeast resulted in 0-1 colonies on average. Importantly, of the WT *DNL4* isolates, 73% had properly excised the entire cassette whereas this was found to be 0% for *dnl4Δ* yeast across numerous independent trials (Fig. 3C). The partial loss of function *dnl4* alleles provided a range of NHEJ efficiencies: the K742A mutant averaged 4 colonies/trial with an excision rate of nearly 50% and other substitutions displayed excision rates of between 0-25%. As expected, the C-terminal *dnl4* truncation at L750 phenocopied the null allele. Diagnostic PCRs confirmed the presence or absence of the Cas9-Kan^R expression cassette for clonal isolates from each of the aforementioned haploid strains tested (Fig. 3D). For strains that had undergone editing and marker excision, the *HIS3* locus was amplified and sequenced; NHEJ followed by imprecise ligation introduced either insertions or deletions at the site of Cas9 cleavage 3 bp upstream of the 5' end of the PAM sequence (Fig. 3E). Finally, each of the *dnl4* alleles was tested within our MGD system as a native cargo-based delivery system (Fig. 3F, *top*). Given that our artificial *DNL4* target was the *dnl4ΔKan^R* null allele, we recognize that in the context of a [gene drive x WT] diploid genome, further modifications would be required to bias the HR-based repair of the intended *dnl4* allele. This could include recoding (silent substitutions) of the *DNL4* C-terminal domain sequence to prevent promiscuous cross-over downstream of the intended mutation(s). Following expression of Cas9

18

and activation of the MGD, the growth profiles of 7 diploid strains were assessed in triplicate and demonstrated efficient drive activity at all three loci (Fig. 3F, *bottom*). Moreover, PCRs from clonal isolates confirmed the presence or absence of each drive and target locus (Supplementary Fig. 6). These data demonstrate that the MGD strategy can be used as a knock-out or allele replacement strategy (at a native locus) with only minimal added sequence (782 bp).

DISCUSSION

In this study, we have developed a multi-locus CRISPR gene drive with a minimal design (MGD) that allows for multiplexing of Cas9 *in trans* across three distinct chromosomal locations (Fig. 4). An alternative strategy could also be employed to create more than one gene drive system within a single genome—a CGD where each locus of interest contains the full complement of genetic information (nuclease, UTR, sgRNA, and optional cargo). In this way, each drive would be completely independent from all other drive(s). While this design clearly provides a maximum level of potential redundancy, there are other technical and safety issues inherent to this multi-nuclease arrangement. For one, countering or inhibiting a CGD with more than one active nuclease would require more sophisticated anti-drive systems, the discovery of additional anti-CRISPR proteins, or complex regulatory systems to ensure inactivation or destruction of each drive. In contrast, our minimal GD design can be inhibited by the AcrIIA2/A4 proteins (BASGALL *et al.* 2018), self-excised by our flanking (u2) sites (ROGGENKAMP *et al.* 2018), or targeted by an anti-drive system no different than a traditional single-locus gene drive. We argue that this type of design provides a higher level of biosecurity and can still accomplish the same task as n-number of "full" gene drives. Moreover, the issue of tightly regulated control of the nuclease transcript may pose additional challenges if the same promoter elements are positioned across multiple chromosomes and epigenetic landscapes in the CGD design.
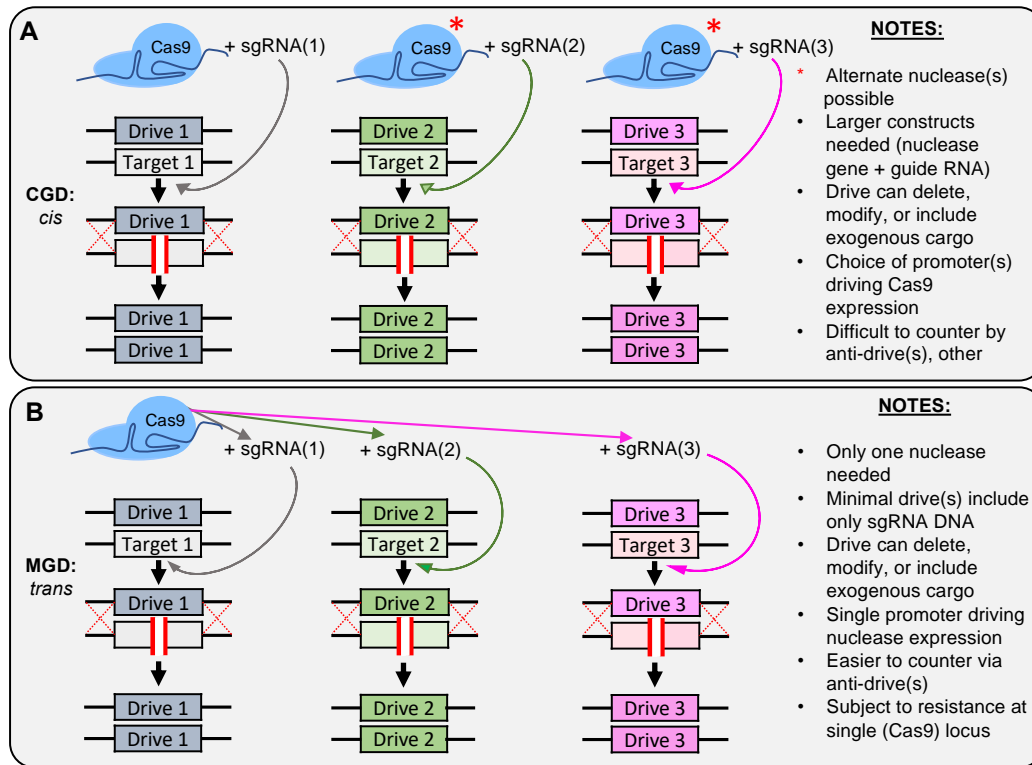
19

Figure 4. Models for multi-locus CRISPR gene drive systems. (A) A proposed gene drive arrangement *in cis*. Each locus to be modified contains a "complete" system (drive and guide RNA cassette). These may be identical nuclease genes, altered variants, or sourced from separate species (e.g. Cas9 versus Cas12a). The action of each drive is fully independent from other drive-containing loci. (B) A single nuclease functions *in trans* across multiple loci with separate guide RNAs. This "minimal" design allows for greater safety and security (easily countered by a single anti-drive system or other means) but may be more susceptible to resistance at the primary (Cas9-harboring) locus.

One potential issue facing our MGD design (or *any* gene drive design, for that matter) is that of natural or evolved resistance to the action of the drive. Since our multi-locus arrangement includes only a single nuclease gene powering all drives, any resistance or escaped action to the

Cas9-containing locus would render all three gene drives inactivated in subsequent generations. However, evidence now exists both *in silico* (MARSHALL *et al.* 2017; NOBLE *et al.* 2017; PROWSE *et al.* 2017) and *in vivo* (CHAMPER *et al.* 2018) that the addition of multiple guide RNAs (to the same genetic target) can reduce (or potentially *eliminate*) resistance to the gene drive. Using current estimations for a given target, five separate guide RNAs may provide a sufficiently rare or improbable event requiring mismatch or mutation to occur at all five target DNA sites (MARSHALL *et al.* 2017; NOBLE *et al.* 2017). This would provide greater than 99% confidence in eliminating an *A. gambiae* population on a continent-wide scale (MARSHALL *et al.* 2017). Therefore, our recommendation would be to greatly bias multiplexing (via multiple guide RNAs) to the gene drive locus harboring the sole copy of Cas9 in the MGD design (in our system, Cas9 creates two DSBs flanking the entire locus). To ensure even higher fidelity of the nuclease and to combat resistance, one could combine the two strategies (CGD and MGD) to have a secondary copy of the nuclease (of the same variant or a different species) positioned at a second locus—additional multiplexing across numerous other loci could include the minimal design (guide RNA cassette only). Finally, while our MGD methodology includes a gene drive consisting of only 455 nucleotides (sgRNA expression cassette), this could be reduced even further to be only a few bases or the absence of any base pairs. Additional sgRNA cassettes could be installed at one or more loci to allow for targeting of chromosomal positions where the "drive" is nothing more than a single base substitution or deletion. Provided few bases separate the DSB site and the intended mutation(s), HR-based repair would allow for propagation of the few bases no different than a "full" gene drive (consisting of many thousands or tens of thousands of bases) into the homozygous condition. The only requirement would involve the sgRNA expression cassette(s) to also be installed within a drive-containing locus *in trans*.

Finally, we chose one of the chromosomal targets within our minimal gene drive system to include both deletions, truncations, and substitution alleles of *DNL4*—one of the essential components of the NHEJ repair pathway. While loss of DNA Ligase IV is non-lethal in yeast (WILSON *et al.* 1997) and flies (GORSKI *et al.* 2003), it is embryonic lethal in mouse (BARNES *et al.* 1998) and not tolerated in mosquito (BASU *et al.* 2015). However, suppression or inhibition of this enzyme or the NHEJ repair pathway has been shown to increase rates of recombination and genomic integration of exogenous DNA using CRISPR systems *in vivo* (BASU *et al.* 2015; CHU *et al.* 2015; MARUYAMA *et al.* 2015; ROBERT *et al.* 2015; VARTAK AND RAGHAVAN 2015; CEN *et al.* 2017; CANNY *et al.* 2018). We demonstrate that this critical factor could be an additional target for a multi-locus gene drive system—suppression of NHEJ, whether by mutated alleles, regulation of transcript, or direct inhibition of the enzyme—would aid in successful HR-based copying of the drive and further reduce the possibility for drive resistance, especially when coupled with multiple guide RNAs. Numerous strategies might be employed to accomplish targeted suppression of NHEJ including testing of additional Ligase IV loss of function alleles that may be widely conserved across eukaryotes; our study focused on the C-terminal BRCT-domain containing portion of Dnl4, but other substitutions have been also been characterized within the N-terminal catalytic domain (CHIRUVELLA *et al.* 2013a).

A multi-locus CRISPR gene drive system should help advance current designs and provide additional options for (i) biosecurity, (ii) drive redundancy, (iii) combatting of evolved resistance, (iv) native gene replacement, (v) multiple gene cargo/genetic pathway delivery, (vi) suppression of NHEJ or activation of HR-promoting repair pathways, and (vii) multiple phenotypic outcomes. Advanced drive arrangements (DHOLE *et al.* 2018) could accomplish multiple outcomes within a single-genome system—the additional of exogenous cargo could also be paired with (native) allele

22

introduction *and* modulating of organism fitness by perturbing numerous other genetic pathways in a single step. As the design and application of CRISPR gene drives continues to advance, we continues to stress the need for multiple levels of control, tunability, inhibition, and drive reversal.

METHODS

Yeast Strains and Plasmids

Standard molecular biology protocols were used to engineer all *S. cerevisiae* strains (Supplementary Table 1) used in this study (SAMBROOK AND RUSSELL 2001). The overall methodology for construction of the triple gene drive strain utilized both standard HR-based chromosomal integrations (*sans* any DSB) and Cas9-based editing (Supplementary Fig. 1). Briefly, DNA constructs were first assembled onto *CEN*-based plasmids (typically pRS315) using *in vivo* assembly in yeast (FINNIGAN AND THORNER 2015). If necessary, point mutations were introduced using PCR mutagenesis (ZHENG *et al.* 2004). Next, the engineered cassette was amplified with a high-fidelity polymerase (KOD Hot Start, EMD Millipore), transformed into yeast using a modified lithium acetate method (ECKERT-BOULET *et al.* 2012), and integrated at the desired genomic locus. PCR was used to diagnose proper chromosomal position for each integration event followed by DNA sequencing. The DNA maps for manipulated yeast loci are included in Supplementary Fig. 2. DNA plasmids used in this study can be found in Supplementary Table 2. Expression cassettes for sgRNA were based on a previous study (DICARLO *et al.* 2013), purchased as synthetic genes (Genscript), and sub-cloned to high-copy plasmids using unique flanking restriction sites. All vectors were confirmed by Sanger sequencing.

Culture Conditions

Budding yeast were cultured in liquid or solid medium. YPD-based medium included 2% peptone, 1% yeast extract, and 2% dextrose. Synthetic (drop-out) medium included yeast nitrogen

base, ammonium sulfate, and amino acid supplements. A raffinose/sucrose mixture (2%/0.2%) was used to pre-induce cultures prior to treatment with galactose (2%). All media was autoclaved or filter sterilized (sugars).

Cas9-based editing *in vivo*

Editing of haploid *S. cerevisiae* strains was performed as previously described (ROGGENKAMP *et al.* 2018). Briefly, an integrated copy of *S. pyogenes* Cas9 was designed with two flanking "unique" (u2) sites—23 base pairs artificially introduced into the genome. This sequence contains a maximum mismatch to the native yeast genome and is used in order to (i) multiplex at two separate sites using a single guide RNA construct, (ii) minimize (or likely eliminate) potential off-target effects, and (iii) allow for increased biosecurity in testing of active CRISPR gene drive systems (FINNIGAN AND THORNER 2016). Haploid yeast were pre-induced overnight in a raffinose/sucrose mixture to saturation, back-diluted to an OD_{600} of approximately 0.35 in rich medium containing galactose, and cultured for 4.5 hr at 30°C. Equimolar amounts (1,000 ng) of high-copy plasmid (sgRNA) were transformed into yeast followed by recovery overnight in galactose and a final plating onto SD-LEU medium. Colonies were imaged and quantified after 3-4 days of growth. Haploid yeast editing experiments included three replicates in triplicate—all as separate transformation events—for each strain (n=9).

Gene drive activation and containment

Haploid yeast strains harboring the gene drive (Cas9) system were first transformed with the sgRNA-containing plasmid (*LEU2*-marked). Next, drive strains were mated to target strains of the opposite mating type on rich medium for 24 hr. Third, yeast were velvet-transferred to synthetic drop-out medium to select diploids (e.g. SD-URA-LEU or SD-URA-LEU-HIS); each haploid genome contained at least one unique selectable marker. Diploids were selected three consecutive

24

rounds with 1-2 days incubation at each step. Fourth, yeast were cultured in pre-induction medium (raffinose/sucrose) lacking leucine overnight, back-diluted into rich medium containing galactose, and grown for 5 hr (or appropriate time intervals). Strains were diluted to approximately 100-500 cells per mL and plated onto SD-LEU for 2 days. Finally, colonies were transferred to the appropriate selection plates (e.g. SD-HIS, G418, SD-URA, and a fresh SD-LEU plate) for 1 additional day of growth before being imaged. The number of surviving colonies on each media type was quantified; experiments were performed in at least triplicate.

A number of safeguards were included in the design of all gene drive systems. First, the genomic targets for all guide RNAs included only non-yeast sequences (u1, GFP, and Kan^R) (FINNIGAN AND THORNER 2016; ROGGENKAMP *et al.* 2017). Second, the primary guide RNA cassette (u1) for targeting of the *HIS3* locus which included Cas9 was maintained on an unstable high-copy (2μ) plasmid; previous work has demonstrated loss of this vector type in the absence of active selection (DICARLO *et al.* 2015; ROGGENKAMP *et al.* 2018). Third, the *S. cerevisiae* BY4741/BY4742 genetic background does not readily undergo sporulation, even under optimal conditions. Fourth, Cas9 expression was repressed by growth on dextrose until gene drives were activated. And finally, all diploid strains, plates, and consumables were autoclaved and inactivated.

Graphics and Evolutionary Analysis

Molecular graphics were generated using the Chimera software package from the Univ. of California, San Francisco (PETTERSEN *et al.* 2004). Homologous sequences to the yeast DNA Ligase IV (Dnl4) protein were obtained using multiple BLAST (NCBI) searches within either the fungal or metazoan clade (Supplementary Table 3). The phylogenetic tree of DNA Ligase IV was created using the Phylogeny.fr software (DEREEPER *et al.* 2008; DEREEPER *et al.* 2010). Multiple sequence alignments were performed using Clustal Omega (LI *et al.* 2015). The predicted

structures of the human, yeast, and mosquito Ligase IV enzyme were generated using I-TASSER (ROY *et al.* 2010). The template structures included the human Lig4 N-terminus (PDB:3W1B) (OCHI *et al.* 2013) and the yeast Dnl4 C-terminus (PDB:1Z56) (DORE *et al.* 2006). Predicted models were individually aligned against the crystal structures using MatchMaker in Chimera. Metrics for the predicted structures are included in Supplementary Table 4.

ACKNOWLEDEGEMENTS & FUNDING

CONFLICT OF INTEREST STATEMENT

G.C.F. (Kansas State University) has filed for a provisional patent entitled "Multi-Locus Gene Drive System," U.S. Serial No. 62/697,855 on July 13, 2018 for the intellectual property and technology described within this work.

ANIMAL AND HUMAN SUBJECT STATEMENT

This study does not use any animals or human subjects.

REFERENCES

Altmann, T., and A. R. Gennery, 2016 DNA ligase IV syndrome; a review. Orphanet J Rare Dis 11: 137.

Barnes, D. E., G. Stamp, I. Rosewell, A. Denzel and T. Lindahl, 1998 Targeted disruption of the gene encoding DNA ligase IV leads to lethality in embryonic mice. Curr Biol 8: 1395-1398.

Basgall, E. M., S. C. Goetting, M. E. Goeckel, R. M. Giersch, E. Roggenkamp *et al.*, 2018 Gene drive inhibition by the anti-CRISPR proteins AcrIIA2 and AcrIIA4 in Saccharomyces cerevisiae. Microbiology 164: 464-474.

Basu, S., A. Aryan, J. M. Overcash, G. H. Samuel, M. A. Anderson *et al.*, 2015 Silencing of end-joining repair for efficient site-specific gene insertion after TALEN/CRISPR mutagenesis in Aedes aegypti. Proc Natl Acad Sci U S A 112: 4038-4043.

Buchman, A., J. M. Marshall, D. Ostrovski, T. Yang and O. S. Akbari, 2018 Synthetically engineered Medea gene drive system in the worldwide crop pest Drosophila suzukii. Proc Natl Acad Sci U S A 115: 4725-4730.

Bull, J. J., and J. E. Barrick, 2017 Arresting Evolution. Trends Genet 33: 910-920.

Bull, J. J., and H. S. Malik, 2017 The gene drive bubble: New realities. PLoS Genet 13: e1006850.

Burt, A., and A. Crisanti, 2018 Gene Drive: Evolved and Synthetic. ACS Chem Biol 13: 343-346.

Canny, M. D., N. Moatti, L. C. K. Wan, A. Fradet-Turcotte, D. Krasner *et al.*, 2018 Inhibition of 53BP1 favors homology-dependent DNA repair and increases CRISPR-Cas9 genome-editing efficiency. Nat Biotechnol 36: 95-102.

Cen, Y., B. Timmermans, B. Souffriau, J. M. Thevelein and P. Van Dijck, 2017 Comparison of genome engineering using the CRISPR-Cas9 system in C. glabrata wild-type and lig4 strains. Fungal Genet Biol 107**:** 44-50.

Champer, J., J. Liu, S. Y. Oh, R. Reeves, A. Luthra *et al.*, 2018 Reducing resistance allele formation in CRISPR gene drive. Proc Natl Acad Sci U S A.

Champer, J., R. Reeves, S. Y. Oh, C. Liu, J. Liu *et al.*, 2017 Novel CRISPR/Cas9 gene drive constructs reveal insights into mechanisms of resistance allele formation and drive efficiency in genetically diverse populations. PLoS Genet 13**:** e1006796.

Chiruvella, K. K., Z. Liang, S. R. Birkeland, V. Basrur and T. E. Wilson, 2013a Saccharomyces cerevisiae DNA ligase IV supports imprecise end joining independently of its catalytic activity. PLoS Genet 9**:** e1003599.

Chiruvella, K. K., Z. Liang and T. E. Wilson, 2013b Repair of double-strand breaks by end joining. Cold Spring Harb Perspect Biol 5**:** a012757.

Chiruvella, K. K., B. M. Renard, S. R. Birkeland, S. Sunder, Z. Liang *et al.*, 2014 Yeast DNA ligase IV mutations reveal a nonhomologous end joining function of BRCT1 distinct from XRCC4/Lif1 binding. DNA Repair (Amst) 24**:** 37-45.

Chistiakov, D. A., 2010 Ligase IV syndrome. Adv Exp Med Biol 685**:** 175-185.

Chu, V. T., T. Weber, B. Wefers, W. Wurst, S. Sander *et al.*, 2015 Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. Nat Biotechnol 33**:** 543-548.

Courtier-Orgogozo, V., B. Morizot and C. Boete, 2017 Using CRISPR-based gene drive for agriculture pest control. EMBO Rep.

Dereeper, A., S. Audic, J. M. Claverie and G. Blanc, 2010 BLAST-EXPLORER helps you building datasets for phylogenetic analysis. BMC Evol Biol 10**:** 8.

Dereeper, A., V. Guignon, G. Blanc, S. Audic, S. Buffet *et al.*, 2008 Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 36**:** W465-469.

Dhole, S., M. R. Vella, A. L. Lloyd and F. Gould, 2018 Invasion and migration of spatially self-limiting gene drives: A comparative analysis. Evol Appl 11**:** 794-808.

DiCarlo, J. E., A. Chavez, S. L. Dietz, K. M. Esvelt and G. M. Church, 2015 Safeguarding CRISPR-Cas9 gene drives in yeast. Nat Biotechnol 33**:** 1250-1255.

DiCarlo, J. E., J. E. Norville, P. Mali, X. Rios, J. Aach *et al.*, 2013 Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res 41**:** 4336-4343.

Dore, A. S., N. Furnham, O. R. Davies, B. L. Sibanda, D. Y. Chirgadze *et al.*, 2006 Structure of an Xrcc4-DNA ligase IV yeast ortholog complex reveals a novel BRCT interaction mode. DNA Repair (Amst) 5**:** 362-368.

Doudna, J. A., and E. Charpentier, 2014 Genome editing. The new frontier of genome engineering with CRISPR-Cas9. Science 346**:** 1258096.

Drury, D. W., A. L. Dapper, D. J. Siniard, G. E. Zentner and M. J. Wade, 2017 CRISPR/Cas9 gene drives in genetically variable and nonrandomly mating wild populations. Sci Adv 3**:** e1601910.

Eckert-Boulet, N., M. L. Pedersen, B. O. Krogh and M. Lisby, 2012 Optimization of ordered plasmid assembly by gap repair in Saccharomyces cerevisiae. Yeast 29**:** 323-334.

Ellenberger, T., and A. E. Tomkinson, 2008 Eukaryotic DNA ligases: structural and functional insights. Annu Rev Biochem 77**:** 313-338.

Esvelt, K. M., and N. J. Gemmell, 2017 Conservation demands safe gene drive. PLoS Biol 15**:** e2003850.

Finnigan, G. C., and J. Thorner, 2015 Complex *in vivo* ligation using homologous recombination and high-efficiency plasmid rescue from *Saccharomyces cerevisiae*. Bio-protocol 5**:** e1521. http://www.bio-protocol.org/e1521

Finnigan, G. C., and J. Thorner, 2016 mCAL: a new approach for versatile multiplex action of Cas9 using one sgRNA and loci flanked by a programmed target sequence. G3 (Bethesda) 6**:** 2147-2156.

Gantz, V. M., N. Jasinskiene, O. Tatarenkova, A. Fazekas, V. M. Macias *et al.*, 2015 Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito Anopheles stephensi. Proc Natl Acad Sci U S A 112**:** E6736-6743.

Gaudelli, N. M., A. C. Komor, H. A. Rees, M. S. Packer, A. H. Badran *et al.*, 2017 Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. Nature 551**:** 464-471.

Godfray, H. C. J., A. North and A. Burt, 2017 How driving endonuclease genes can be used to combat pests and disease vectors. BMC Biol 15**:** 81.

Gorski, M. M., J. C. Eeken, A. W. de Jong, I. Klink, M. Loos *et al.*, 2003 The Drosophila melanogaster DNA Ligase IV gene plays a crucial role in the repair of radiation-induced DNA double-strand breaks and acts synergistically with Rad54. Genetics 165**:** 1929-1941.

Grunwald, H. A., V. M. Gantz, G. Poplawski, X.-r. S. Xu, E. Bier *et al.*, 2018 Super-Mendelian inheritance mediated by CRISPR/Cas9 in the female mouse germline. bioRxiv.

Hammond, A., R. Galizi, K. Kyrou, A. Simoni, C. Siniscalchi *et al.*, 2016 A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector Anopheles gambiae. Nat Biotechnol 34**:** 78-83.

Hammond, A. M., and R. Galizi, 2017 Gene drives to fight malaria: current state and future directions. Pathog Glob Health 111**:** 412-423.

Hammond, A. M., K. Kyrou, M. Bruttini, A. North, R. Galizi *et al.*, 2017 The creation and selection of mutations resistant to a gene drive over multiple generations in the malaria mosquito. PLoS Genet 13**:** e1007039.

James, S., F. H. Collins, P. A. Welkhoff, C. Emerson, H. C. J. Godfray *et al.*, 2018 Pathway to Deployment of Gene Drive Mosquitoes as a Potential Biocontrol Tool for Elimination of Malaria in Sub-Saharan Africa: Recommendations of a Scientific Working Group(dagger). Am J Trop Med Hyg 98**:** 1-49.

Jinek, M., K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna *et al.*, 2012 A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337**:** 816-821.

Jinek, M., A. East, A. Cheng, S. Lin, E. Ma *et al.*, 2013 RNA-programmed genome editing in human cells. Elife 2**:** e00471.

KaramiNejadRanjbar, M., K. N. Eckermann, H. M. M. Ahmed, C. H. Sanchez, S. Dippel *et al.*, 2018 Consequences of resistance evolution in a Cas9-based sex conversion-suppression gene drive for insect pest management. Proc Natl Acad Sci U S A 115**:** 6189-6194.

Lambert, B., A. North, A. Burt and H. C. J. Godfray, 2018 The use of driving endonuclease genes to suppress mosquito vectors of malaria in temporally variable environments. Malar J 17**:** 154.

Li, W., A. Cowley, M. Uludag, T. Gur, H. McWilliam *et al.*, 2015 The EMBL-EBI bioinformatics web and programmatic tools framework. Nucleic Acids Res 43**:** W580-584.

Lieber, M. R., 2010 The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem 79**:** 181-211.

Marshall, J. M., A. Buchman, C. H. Sanchez and O. S. Akbari, 2017 Overcoming evolved resistance to population-suppressing homing-based gene drives. Sci Rep 7**:** 3776.

Maruyama, T., S. K. Dougan, M. C. Truttmann, A. M. Bilate, J. R. Ingram *et al.*, 2015 Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. Nat Biotechnol 33**:** 538-542.

Misof, B., S. Liu, K. Meusemann, R. S. Peters, A. Donath *et al.*, 2014 Phylogenomics resolves the timing and pattern of insect evolution. Science 346**:** 763-767.

Najjar, D. A., A. M. Normandin, E. A. Strait and K. M. Esvelt, 2017 Driving towards ecotechnologies. Pathog Glob Health 111**:** 448-458.

Noble, C., B. Adlam, G. M. Church, K. M. Esvelt and M. A. Nowak, 2018 Current CRISPR gene drive systems are likely to be highly invasive in wild populations. Elife 7.

Noble, C., J. Olejarz, K. M. Esvelt, G. M. Church and M. A. Nowak, 2017 Evolutionary dynamics of CRISPR gene drives. Sci Adv 3**:** e1601964.

Ochi, T., X. Gu and T. L. Blundell, 2013 Structure of the catalytic region of DNA ligase IV in complex with an Artemis fragment sheds light on double-strand break repair. Structure 21**:** 672-679.

Pettersen, E. F., T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt *et al.*, 2004 UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25: 1605-1612.

Prowse, T. A. A., P. Cassey, J. V. Ross, C. Pfitzner, T. A. Wittmann *et al.*, 2017 Dodging silver bullets: good CRISPR gene-drive design is critical for eradicating exotic vertebrates. Proc Biol Sci 284.

Qi, L. S., M. H. Larson, L. A. Gilbert, J. A. Doudna, J. S. Weissman *et al.*, 2013 Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152: 1173-1183.

Robert, F., M. Barbeau, S. Ethier, J. Dostie and J. Pelletier, 2015 Pharmacological inhibition of DNA-PK stimulates Cas9-mediated genome editing. Genome Med 7: 93.

Roggenkamp, E., R. M. Giersch, M. N. Schrock, E. Turnquist, M. Halloran *et al.*, 2018 Tuning CRISPR-Cas9 Gene Drives in Saccharomyces cerevisiae. G3 (Bethesda) 8: 999-1018.

Roggenkamp, E., R. M. Giersch, E. Wedeman, M. Eaton, E. Turnquist *et al.*, 2017 CRISPR-UnLOCK: multipurpose Cas9-based strategies for Conversion of yeast libraries and strains. Front Microbiol 8: 1773.

Roy, A., A. Kucukural and Y. Zhang, 2010 I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 5: 725-738.

Sambrook, J., and D. W. Russell, 2001 *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schwartz, C., K. Frogue, A. Ramesh, J. Misa and I. Wheeldon, 2017 CRISPRi repression of nonhomologous end-joining for enhanced genome engineering via homologous recombination in Yarrowia lipolytica. Biotechnol Bioeng 114: 2896-2906.

Shapiro, R. S., A. Chavez, C. B. M. Porter, M. Hamblin, C. S. Kaas *et al.*, 2018 A CRISPR-Cas9-based gene drive platform for genetic interaction analysis in Candida albicans. Nat Microbiol 3**:** 73-82.

Thakore, P. I., A. M. D'Ippolito, L. Song, A. Safi, N. K. Shivakumar *et al.*, 2015 Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. Nat Methods 12**:** 1143-1149.

Unckless, R. L., A. G. Clark and P. W. Messer, 2017 Evolution of Resistance Against CRISPR/Cas9 Gene Drive. Genetics 205**:** 827-841.

Vartak, S. V., and S. C. Raghavan, 2015 Inhibition of nonhomologous end joining to increase the specificity of CRISPR/Cas9 genome editing. Febs j 282**:** 4289-4294.

Vella, M. R., C. E. Gunning, A. L. Lloyd and F. Gould, 2017 Evaluating strategies for reversing CRISPR-Cas9 gene drives. Sci Rep 7**:** 11038.

Wilson, T. E., U. Grawunder and M. R. Lieber, 1997 Yeast DNA ligase IV mediates non-homologous DNA end joining. Nature 388**:** 495-498.

Zetsche, B., J. S. Gootenberg, O. O. Abudayyeh, I. M. Slaymaker, K. S. Makarova *et al.*, 2015 Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell 163**:** 759-771.

Zheng, L., U. Baumann and J. L. Reymond, 2004 An efficient one-step site-directed and site-saturation mutagenesis protocol. Nucleic Acids Res 32**:** e115.