

Expanding the CRISPR/Cas9 toolkit for *Pichia pastoris* with efficient donor integration and alternative resistance markers

Astrid Weninger¹ | Jasmin E. Fischer² | Hana Raschmanová³ | Claudia Kniely¹ |
Thomas Vogl¹  | Anton Glieder^{1,2}

¹ Institute of Molecular Biotechnology, Graz University of Technology, Graz, Austria

² Bisy e.U., Wetzawinkel, Hofstätten/Raab, Austria

³ Department of Biotechnology, University of Chemistry and Technology Prague, Prague, Czech Republic

Correspondence

Thomas Vogl, Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel.

Email: thomas.vogl@aon.at

Funding information

European Union project “CHEM21”, Grant number: 115360; European Union project “ROBOX”, Grant number: 635734; Austrian Science Fund (FWF) project number W901 (DK “Molecular Enzymology” Graz)

Abstract

Komagataella phaffii (syn. *Pichia pastoris*) is one of the most commonly used host systems for recombinant protein expression. Achieving targeted genetic modifications had been hindered by low frequencies of homologous recombination (HR). Recently, a CRISPR/Cas9 genome editing system has been implemented for *P. pastoris* enabling gene knockouts based on indels (insertion, deletions) via non-homologous end joining (NHEJ) at near 100% efficiency. However, specifically integrating homologous donor cassettes via HR for replacement studies had proven difficult resulting at most in ~20% correct integration using CRISPR/Cas9. Here, we demonstrate the CRISPR/Cas9 mediated integration of markerless donor cassettes at efficiencies approaching 100% using a *ku70* deletion strain. The Ku70p is involved in NHEJ repair and lack of the protein appears to favor repair via HR near exclusively. While the absolute number of transformants in the $\Delta ku70$ strain is reduced, virtually all surviving transformants showed correct integration. In the wildtype strain, markerless donor cassette integration was also improved up to 25-fold by placing an autonomously replicating sequence (ARS) on the donor cassette. Alternative strategies for improving donor cassette integration using a Cas9 nickase variant or reducing off targeting associated toxicity using a high fidelity Cas9 variant were so far not successful in our hands in *P. pastoris*. Furthermore we provide Cas9/gRNA expression plasmids with a Geneticin resistance marker which proved to be versatile tools for marker recycling. The reported CRISPR-Cas9 tools can be applied for modifying existing production strains and also pave the way for markerless whole genome modification studies in *P. pastoris*.

KEYWORDS

CRISPR/Cas9, homologous recombination, *KU70*, marker recycling, *Pichia pastoris*, resistance markers

Astrid Weninger and Jasmin Fischer contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2017 The Authors. *Journal of Cellular Biochemistry* Published by Wiley Periodicals, Inc.

1 | INTRODUCTION

Within a few years, genome engineering by CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats/CRISPR associated protein 9) became one of the most valuable tools in biosciences.^{1–3} This method facilitates targeted, programmable genome modifications and provides major advantages over classical knock-out-, knock-in-approaches as well as alternative genome engineering strategies. Besides its potency in basic research for studying disease conditions,^{4–7} CRISPR/Cas9 was also used for host (strain) engineering for biotechnological production processes.^{8–10} Prior to CRISPR/Cas9, highly efficient genome manipulations were difficult to achieve in many organisms. Gene targeting with conventional knock-out/knock-in cassettes containing homologous sequences depended on the organisms amenability for homologous recombination and homology cassettes were predominantly integrated ectopically.^{11,12} In addition, the replacement of DNA regions or the insertion of markers (or other disruption cassettes) changes the context of the genome sequence and can influence neighboring genes and their expression. This is especially relevant for maintaining the functional fidelity of tightly packed microbial genomes, where adjacent open reading frames are sometimes only separated by very short DNA stretches. Genome engineering methods available prior to the advent of CRISPR/Cas9 such as Zinc-finger nucleases or TALENs^{12–14} required laborious protein engineering to reprogram the targeting locus. CRISPR/Cas9 allows to easily introduce targeted strand breaks in almost any desired genomic locus and can easily be reprogrammed. The DNA endonuclease Cas9 is guided by a short guide RNA (gRNA, or single guide RNA, sgRNA) to introduce a double strand break (DSB) at genomic regions complementary to the gRNA sequence. The strand breaks are sealed by the cell's endogenous repair machinery, allowing the introduction of various genomic modifications.¹⁵ Reprogramming CRISPR/Cas9 and thereby targeting different loci is performed by changing a 20 bp sequence of the gRNA. Besides a growing list of organisms, in which CRISPR/Cas9 systems were successfully applied for genome engineering,^{1,9,16,17} also continuous improvements of the method itself have been reported. Particularly remarkable are high-fidelity Cas9 enzymes (eSpCas9, SpCas9-HF1), which operate almost exclusively at on-target sites.^{18,19} Thereby unwanted off-targeting, one of the major drawbacks of CRISPR/Cas9, could be reduced.

Due to its potent secretion of pure protein, less extensive glycosylation and high cell density growth *P. pastoris* (*syn. Komagataella phaffii*) is one of the most commonly used eukaryotic expression hosts.^{20,21} While molecular mechanisms in the “classic” yeast *S. cerevisiae* are well-studied, similar information is widely lacking in *P. pastoris*.²² A key

advantage of *S. cerevisiae* has been the ease to perform genome modifications due to its remarkably efficient homologous recombination machinery. For *P. pastoris*, and also other non-conventional yeasts, the deletion of genes and the resulting studies of gene functions are more challenging. Deletion cassettes for targeted gene disruption are often randomly inserted in the genome by ectopic integration (with correct targeting rates of <1–30%²³), requiring laborious screening procedures to identify correct knockout strains. In order to increase the homologous recombination (HR) frequencies in *P. pastoris* for classical HR experiments, a $\Delta ku70$ was generated previously, which lacks the gene coding for the Ku70 protein, a key player in the NHEJ mechanism.²³ Using this strain, targeting efficiencies over 90% with knockout cassettes containing only 250 bp flanking homology arms had been obtained under selective conditions.²³

The transformation of selection marker-free fragments has yet not even been performed with the *P. pastoris* $\Delta ku70$ strain, because it would require intensive screening (since lack of selection after a transformation would result in outgrowth of unmodified cells that did not take up exogenous DNA). However, double HR of a (marker-free) donor DNA cassette is required to completely delete a CDS or to replace it with a different sequence at a genomic locus. Targeted single and double strand break induced DNA repair can be used to increase the HR frequency by several orders of magnitude.^{24–26} When co-transforming homology donor cassettes and CRISPR/Cas9 expression vectors an increase in the HR rate was observed for most yeast species, for which a CRISPR/Cas9 system has been developed.^{27–31} The increase in HR frequency due to CRISPR/Cas9 targeted cleavage also allowed the integration of marker-less donor fragments.

A highly efficient CRISPR/Cas9 system was also implemented for the methylotrophic yeast *P. pastoris*.³² A human codon optimized *CAS9* nuclease and ribozyme-flanked gRNAs were expressed under the control of a bidirectional, constitutive RNA polymerase II promoter (Figure 1A). The ribozymes were required for the processing and functional production of gRNAs without additional 5'/3' sequences (stemming from the 5'/3' UTRs of the promoter/terminators used). The *CAS9/gRNA* expression cassettes were provided on an episomal plasmid and were cured by growth on non-selective media. In *P. pastoris* the CRISPR/Cas9 induced double strand breaks are repaired by the cellular NHEJ machinery, resulting in the formation of short insertions and deletions (indels) in the coding sequence (CDS), and frameshift mutations. The *P. pastoris* CRISPR/Cas9 system enables gene deletions at targeting rates approaching 100% and also the simultaneous deletion of different genes by using multiple gRNAs.³² However, specifically integrating donor cassettes via homologous

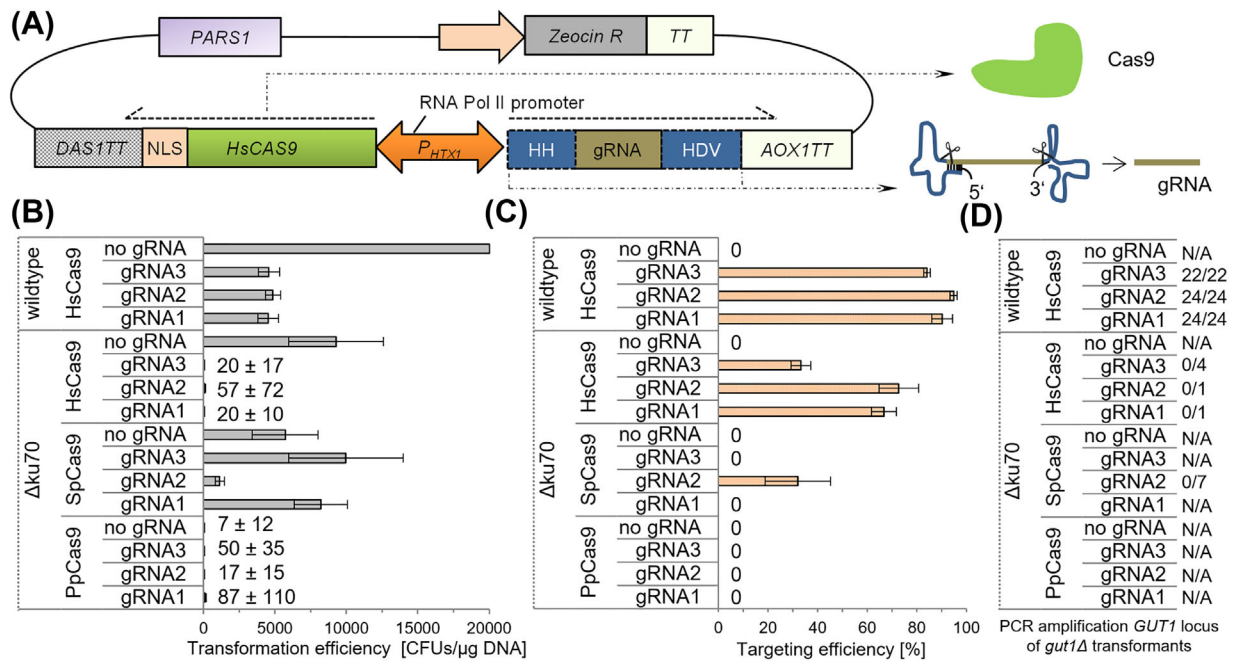


FIGURE 1 Functional Cas9/gRNA expression drastically decreases transformation efficiencies in *P. pastoris* $\Delta ku70$ compared to the wildtype strain. A, Schematic illustration of the Cas9 and gRNA expression strategy successfully calibrated previously³² and used as the basis for this study. A human codon optimized CAS9 gene fused to a C-terminal SV40 nuclear localization sequence and the gRNA flanked by hammerhead (HH) and hepatitis delta virus (HDV) ribozymes are expressed from an episomal plasmid (carrying the PARS1 autonomously replicating sequence and a Zeocin resistance marker). The ribozymes are transcribed and auto-catalytically cleave themselves resulting in a gRNA without additional sequences from the 5' or 3' UTR of the promoter/terminator used. B–D, *P. pastoris* CBS 7435 $\Delta ku70$ and wildtype³² strains were transformed with CRISPR/Cas9 plasmids containing either *PpCAS9*, *HsCAS9* or *SpCAS9* codon optimized sequences and gRNAs to target *GUT1*. Mean values and standard deviations of biological triplicates are shown. B, Transformation efficiencies of *P. pastoris* CBS 7435 $\Delta ku70$ and wildtype with CRISPR/Cas9 plasmids. *P. pastoris* CBS 7435 wildtype results have been reported previously.³² HsCas9 (w/o gRNA) yielded >20 000 *P. pastoris* wildtype transformants and partly a cell lawn, suggesting that the nuclease is barely active having no detrimental effects on cell growth. The functional expression of HsCas9 and different gRNA yielded a reduction of the transformation efficiencies in *P. pastoris* wildtype. The CRISPR/Cas9 systems were even more lethal for the *P. pastoris* $\Delta ku70$ strain, where the transformation efficiencies decreased drastically. C, *GUT1* targeting efficiencies of *P. pastoris* CBS 7435 $\Delta ku70$ and wildtype with CRISPR/Cas9 plasmids. The targeting efficiencies were reduced compared to the *P. pastoris* wildtype strain. Among the different Cas9 different codon optimized sequences, HsCas9 was most effective in *P. pastoris* $\Delta ku70$. D, Amplification of the *GUT1* locus upon CRISPR/Cas9 cleavage. NHEJ-mediated indel mutations were obtained in the *P. pastoris* wildtype strain. The indel mutations were verified by sequencing.³² The amplification of the *GUT1* in the *P. pastoris* $\Delta ku70$ strain was not possible with the same primers used for the *P. pastoris* wildtype strain (see S3)

recombination (HR) for replacement studies had proven difficult resulting in low efficiencies in correct integration using CRISPR/Cas9 in *P. pastoris*.

Here, we expanded the existing CRISPR/Cas9 genome engineering system for *P. pastoris* by overcoming these previous limitations. We demonstrate the integration of marker-free donor cassettes at efficiencies approaching 100% using a *ku70* deletion strain previously reported.²³ Alternatively, in *P. pastoris* wildtype strains integration rates for marker-free donor cassettes of 50% were obtained by using stability donor cassettes with autonomously replicating sequences (ARSs). ARSs are typically placed on plasmids to allow episomal expression, which is beneficial for the

production of selected proteins.³² Depending on the ARS multiple copies of the plasmids are present in the cell (PARS1, 10 copies) and the plasmids are maintained in the presence of selective pressure.³³ Another previously reported strategy to increase the HR frequencies is the expression of the Cas9 nickase, which was also evaluated in this study. This Cas9 variant creates single strand breaks due to the inactivation of one of its catalytic domains via the introduction of a point mutation.^{34,35} By introducing single strand breaks an increase in donor recombination was demonstrated by^{36,37} in mammalian cells. However, the use of the Cas9 nickase did so far not show clear advantages in our hands in *P. pastoris*.

2 | MATERIALS AND METHODS

2.1 | Chemicals

Enzymes were obtained from Thermo Fisher Scientific, Vienna, Austria. D(+)-biotin was obtained from Sigma-Aldrich, Vienna, Austria. Difco yeast nitrogen base w/o amino acids, Bacto tryptone and Bacto yeast extract were obtained from Becton Dickinson, Schwechat, Austria. Zeocin was obtained from Thermo Fisher Scientific. Other chemicals were purchased from Carl Roth, Karlsruhe, Germany. Oligonucleotides were ordered from Integrated DNA Technologies, Leuven Belgium, see supplementary Table S1 for the sequences.

2.2 | Strains and constructs

Yeast strains in this study were based on the *P. pastoris* wildtype strain CBS 7435 (identical to NRRL Y-11430 and ATCC 76273). For some experiments a closely related *P. pastoris* BSYBG11 mutS strains (NRRL Y-11430 Δ aox1 derivatives lacking episomal killer plasmids,³⁸ obtained from bisy e.U., Austria) were used. The plasmids used for the expression of the CRISPR/Cas9 constructs are based on the *E.coli/P. pastoris* shuttle vector pPpT4_GAP (Genbank accession number: JQ519692²³). The generation of several CRISPR/Cas9 constructs had been described in the paper of Weninger et al.³² Here we used only the most efficient designs (number 81-83³²). All plasmids were sequenced by Sanger sequencing. Maps of key plasmids generated in this study are provided in supplementary File S2. Transformations and screenings were performed as described by Weninger et al.³²

2.2.1 | Co-transformation of CRISPR/Cas9 plasmids and donor cassettes in *P. pastoris* CBS7435 Δ ku70

The *P. pastoris* CBS7435 Δ ku70 strain reported previously was used.²³ In short, a knock-in and partial deletion into the *KU70* gene had been performed and in addition the start codon had been also modified to cause a frameshift. The correctness of the modification had been verified and the strain comprehensively characterized.^{23,39} The generation of the donor cassettes had been described previously.³² The donor cassette bearing PARS1 was generated as follows: pJET1.2 blunt-GUT1 (CloneJET PCR Cloning Kit, Thermo Fisher Scientific), consisting of the pJET1.2 backbone and the 5' upstream fused to the 3' downstream regions of the *GUT1* locus, was linearized using *Pst*I. PARS1 was amplified from the plasmid pPpT4_HTX1-PARS1-HsCas9³² using the primer pair Pjet-ARS-fw/Pjet-ARS-rv. The PARS1 PCR fragment was cloned into the vector backbone by assembly

cloning.⁴⁰ One hundred nanograms of the purified circular donor cassettes and 1 μ g of the circular or linearized, purified donor cassettes (*Hind*III linearized pJET1.2 blunt-GUT1-PARS1) were used for the *P. pastoris* transformation. The primers GUTout3prR1 and 5prGUT1_fwd were used for the amplification of the *GUT1* locus and the primer PGUTseq2 was used for sequencing. In case of *P. pastoris* Δ ku70 transformants, which displayed a growth defect on glycerol upon the transformation with circular CRISPR/Cas9 plasmids, the primers GUT1-3-fw and GUTout3prR1 were used for the amplification of the *GUT1* locus and the primers GUT1-438-seq-fw and GUT1-525-seq-rv were used for sequencing.

2.2.2 | Constructs for DNA deletion at the DAS1/DAS2 locus

The plasmid pPpT4_HTX1-PARS1-HsCas9³² was linearized with *Not*I. The gRNAs we ordered on synthetic double stranded DNA fragments (gBlocks, Integrated DNA technologies, RZ-DAS1/2-gRNA1-RZ and RZ-DAS1/2-gRNA2-RZ) and cloned into the linearized plasmid by assembly cloning. One hundred nanograms plasmid DNA were used for the *P. pastoris* transformations. Randomly picked transformants were cultivated in 96-deep well plates (96-DWPs, 250 μ L YPD-Zeocin [50 μ g/mL]) and stamped on agar plates with either glucose (BMD1) or methanol (BMM1) as carbon source. The primers Das2_rev and Das1_fw were used for the amplification and for sequencing of the *DAS1/DAS2* locus, the primers Das2_rev and Das2_fw were used for the amplification of the *DAS2* locus, and the primers Das1_rev and Das1_fw were used for the amplification of the *DAS1* locus.

2.2.3 | CRISPR/Cas9 nickase constructs

For the construction of vectors bearing *HsCas9* nuclease, the plasmid pPpT4_HTX1-PARS1-HsCas9³² was linearized with *Not*I. The gRNAs we ordered on synthetic double stranded DNA fragments (gBlocks, Integrated DNA technologies, RZ-GUT1-gRNA0-RZ, and RZ-GUT1-gRNA4-RZ) and cloned into the linearized plasmid by assembly cloning. For the construction of vectors bearing *HsCas9* nickase, the plasmid pPpT4_HTX1-PARS1³² was linearized with *Eco*RI. *HsCas9* was amplified from the plasmid pPpT4_HTX1-PARS1-HsCas9 using the primers DAS1TT-hsCas9-rv and HsCas9-D10A-fw, containing the point mutation 10D \rightarrow A (codon GAT to GCT). The PCR amplicon and the vector backbone were assembled to generate pPpT4_HTX1-PARS1-HsCas9-nickase. pPpT4_HTX1-PARS1-HsCas9-nickase was linearized with *Not*I and the gRNAs GUT1-gRNA0 and GUT1-gRNA4 were cloned into the backbone.

2.2.4 | Construction of CRISPR/Cas9 plasmid bearing high fidelity HsCas9

The vector pPpT4_GAP1-PARS1-Zeo³² was cleaved with *PciI* and *NotI* to remove the *GAP* promoter (*P_{GAP}*). A fragment containing the *DASITT* terminator, the SV40 nuclear localization signal and a part of *HsCas9* coding sequence (bp 2780-4104) was PCR amplified from plasmid pPpT4_pHTX1-HsCas9-RZ-GUT1-gRNA1³² using primers P_AOX1_syn_pUCORI_Gibs_fv and HsCas9_rev. A fragment containing a part of *HsCas9* coding sequence (bp 1-2255), the *P_{HTX1}* bidirectional promoter, the hammerhead ribozyme (HH), GUT1-gRNA1, and the hepatitis delta virus ribozyme (HDV) was also PCR amplified from the plasmid pPpT4_pHTX1-HsCas9-RZ-GUT1-gRNA1 using primers HsCas9_fv and HDV_AOX1TT_Gibs_rev. The part of *HsCas9* coding sequence, in which four amino acids should be exchanged according to Kleinstiver et al¹⁸ in order to prepare the off-target free version of *HsCas9* (497N → A, 661R → A, 695Q → A, 926Q → A), was ordered as synthetic double stranded DNA (gBlock_hf-HsCas9). GCT was used as codon for alanine at all four positions. The two PCR fragments and the synthetic double stranded DNA fragment were cloned in the vector backbone by assembly cloning to form the plasmid pPpT4_pHTX1-hf-HsCas9-GUT1-gRNA1 (supplementary File S2). For the genotypic characterization the *GUT1* locus was amplified from colonies applying the Phire Plant Direct PCR Master Mix (Thermo Fisher Scientific, Vienna, Austria) with primers 3UTRGUTR and seq-pGUT1-332..308-fwd. Primer seq-pGUT1-332..308-fwd was used for sequencing.

2.2.5 | CRISPR/Cas9 plasmids bearing a geneticin resistance cassette

The plasmid pPpT4_HTX1-PARS1-HsCas9³² was cut with the restriction enzymes *XbaI* and *SmiI* to remove the Zeocin resistance marker cassette. Accordingly, the plasmid pPpT4_S_Kan (Genbank accession number: JQ19694²³) was cut with the restriction enzymes *XbaI* and *SmiI* to obtain a fragment containing the Geneticin resistance marker cassette. The backbone and the insert were assembled by ligation with the T4 DNA ligase to create pPpT4_HTX1-PARS1-HsCas9-Gen, supplementary File S2. pPpT4_HTX1-PARS1-HsCas9-Gen was linearized with *NotI* and the gRNAs RZ-GUT1-gRNA1-RZ, RZ-GUT1-gRNA2-RZ, RZ-GUT1-gRNA3-RZ to target *GUT1* as well as the gRNAs ZEO-gRNA1, ZEO-gRNA2, ZEO-gRNA3 to target the Zeocin resistance gene, ordered on synthetic double stranded DNA fragments, were cloned into the plasmid by assembly cloning. The *P. pastoris* CBS7435 wildtype strain expressing enhanced green fluorescent protein (eGFP,⁴¹ Institute of Molecular Biotechnology (IMBT, Graz University of

Technology, Austria) strain collection number: 8008) and the *P. pastoris* BSYBG11 mutS strains expressing glucose oxidase (*GOX*, accession number: AID16306.1, synthetic gene, codon optimized for *P. pastoris*, IMBT strain collection number: 8009), and *Candida antarctica* lipase B (*CALB*,⁴² synthetic gene, codon optimized for *P. pastoris*, IMBT strain collection number: 8010) from ectopically integrated expression cassettes in their genomes were transformed with 100 ng circular CRISPR-Cas9 plasmid DNA. Random transformants were cultivated in 96-DWPs containing 250 μL YPD-Geneticin media (300 μg/mL Geneticin) to support maintenance of ARS plasmids. After 60 h the transformants were stamped on YPD, YPD-Zeocin (50 μg/mL) and YPD-Geneticin (300 μg/mL) agar plates to determine the targeting efficiency by phenotypic characterization.

3 | RESULTS AND DISCUSSION

3.1 | Increasing the efficiency of homology-directed repair (HDR) in *P. pastoris*

3.1.1 | CRISPR-Cas in *P. pastoris* $\Delta ku70$ allows integration of donor cassettes at near 100% efficiency

NHEJ-repair appears strongly impaired using CRISPR/Cas9 in *P. pastoris* $\Delta ku70$ strain

As the CRISPR/Cas9 mediated integration of donor cassettes in *P. pastoris* wildtype strains was so far inefficient and in part even unsuccessful,³² we evaluated to combine the potency of the *P. pastoris* $\Delta ku70$ strain and CRISPR/Cas9 to increase the HR rates and to specifically enable the integration of markerless fragments. Prior to integration studies using homology donor cassettes, we characterized CRISPR/Cas9 mediated gene deletion by the double strand cleavage in *P. pastoris* $\Delta ku70$. Due to the impaired NHEJ machinery of *P. pastoris* $\Delta ku70$ CRISPR/Cas9 expression might be toxic and/or lethal, even if homology donor fragments are co-transformed. The *P. pastoris* wildtype and $\Delta ku70$ strains were transformed with CRISPR/Cas9 plasmids (Cas9 + gRNA) without homology donor cassettes (Figure 1B-D). The CRISPR/Cas9 vectors bore ribozyme flanked gRNAs to target *GUT1* and *CAS9* sequences with different codon optimization (*P. pastoris* codon optimized *CAS9* [*PpCAS9*], *Streptococcus pyogenes* *CAS9* [*SpCAS9*], and a human codon optimized *CAS9* [*HsCAS9*]) under the control of a bidirectional promoter.³² In our previous study in the wildtype strain,³² it appeared that *SpCas9* is poorly expressed, *PpCas9* is expressed at too high levels and only *HsCas9* is giving optimal expression permitting high genome editing efficiencies. Hence, we tested here again all three available Cas9 variants, to evaluate if for the *P. pastoris* $\Delta ku70$ strain a different expression level

would be needed (for the wildtype strain only the established *HsCas9* was included).

The simple *GUT1* locus was targeted as *GUT1* defective strains display a reduced growth phenotype on glycerol as carbon source compared to the wild type strain.²³ Hence mutations in the *GUT1* coding sequence (CDS) could be easily detected by replica-stamping the cells on glycerol plates. The transformation rates of the $\Delta ku70$ strain with the CRISPR/Cas9 plasmids were lower compared to the wildtype strain (*P. pastoris* wildtype: ~5000 colony forming units [CFUs]/ μg DNA, depending on the gRNA used³²). Less than 60 *P. pastoris* $\Delta ku70$ CFUs/ μg plasmid DNA were obtained with the constructs bearing either *HsCas9* or *PpCas9* and one of the three gRNAs (Figure 1B) yielding an almost 100-fold reduction compared to the transformation of wildtype strain (expressing *HsCAS9*). None of the *PpCas9* transformants had growth deficiencies on glycerol plates indicating that no double strand break followed by error prone NHEJ-repair took place at the target locus (Figure 1C). 33-73% of the clones bearing *HsCas9* displayed a reduced growth phenotype on glycerol. The targeting efficiency of *HsCas9* depended on the gRNA used and was reduced compared to the *P. pastoris* wildtype strain (where 87-94% of transformants could not grow on glycerol), confirming the expectation that the *Ku70* protein is involved in the error prone repair of double strand breaks caused by the targeted Cas9 mediated double strand break. The number of clones obtained after the transformation with *SpCas9* differed depending on the gRNA sequence. Transformations with *SpCas9* and gRNA1 or gRNA3 resulted in more than 10 000 CFUs/ μg DNA. None of the transformants showed reduced growth on glycerol. Approximately 1500 CFUs/ μg DNA were obtained, when a plasmid bearing *SpCas9* and gRNA2 had been transformed. Thereof 32% of the transformants had a reduced growth phenotype on glycerol. This result suggests that, in combination with the *SpCAS9* sequence, gRNA1 and gRNA3 are not functional or at least less functional than gRNA2. At the same time, functional gRNA expression appears linked to reduced transformation rates, possibly due to double strand breaks that could not be repaired efficiently in the NHEJ deficient $\Delta ku70$ strain. This notion is supported by the results of *HsCAS9* in the $\Delta ku70$ strain were even higher targeting efficiencies correlated with a higher reduction in transformation rates. These observations were not surprising considering the fact that translocations are common, when proteins involved in NHEJ are missing.⁴³ Also in the absence of gRNAs a reduction in the targeting efficiency was observed for the *P. pastoris* wildtype³² and $\Delta ku70$ strain, when expressing *HsCas9* and *PpCas9*, which indicates a toxic effect due to the abundance of high levels of the Cas9 protein localized in the nucleus (which may interfere with cellular processes such as DNA replication, transcription and transcription factor binding). The $\Delta ku70$ strain

appears to be even more susceptible to this effect (than the wildtype strain) possibly owing to its reduced stress resistance.²³

Genomic DNA of growth deficient and normally growing transformants of all constructs was isolated. Interesting, we were not able to amplify the *GUT1* locus of in $\Delta gut1$ mutants with primers, which had been used for the amplification of genomic DNA isolated from *P. pastoris* wildtype CRISPR/Cas9 transformants (Figure 1D).

These results strongly suggest that NHEJ is practically completely abolished in the $\Delta ku70$ strain: First, the transformation rates were drastically reduced compared to the wildtype strain (~100-fold) when coexpressing gRNAs, indicating that DSBs introduced by Cas9 cannot be repaired and are lethal. Expression of Cas9 without gRNAs in the $\Delta ku70$ strain shows only a moderate decrease in the transformation rate to about half of the wildtype strain. This effect may be attributable to different amenabilities of the $\Delta ku70$ strain to transformations or different tolerance to stress caused from expression of substantial amounts of the heterologous Cas9 protein. Second, none of the *P. pastoris* $\Delta ku70$ $\Delta gut1$ strains obtained tested positive for NHEJ mutations (Figure 1D). While we could PCR amplify the targeted *GUT1* locus in all 70 transformants for the wildtype strain, we could not get PCR products for any of the 11 clones $\Delta ku70$ transformants with the primers used for the amplification of the *GUT1* locus of wildtype transformants. However, when using a forward primer, which bound further upstream of the Cas9 cleavage site, the amplification of the *GUT1* locus was possible (supplementary Figure S3). Sanger sequencing revealed that DNA fragments of approximately 600 bp upstream of the Cas9 cleavage position were deleted in the *P. pastoris* $\Delta ku70$ strain. This correlates to findings in various eukaryotic hosts, where the deletion of *KU* genes resulted in extended DNA degradation and increased deletion sizes upon double strand cleavage.⁴⁴ These results suggest, that most probably the DSB in the $\Delta ku70$ was repaired by other means than NHEJ, for example translocation events taking place, which caused the rearrangement with non-homologous parts of the chromosomes, or by microhomology-mediated end joining (MMEJ).⁴⁵

P. pastoris $\Delta ku70$ transformants with wildtype like growth did not contain a mutation in the target sequence. CRISPR/Cas9 expression has already been described in various NHEJ-deficient organisms, mainly in combination with homology donor fragments.^{28,46,47} Similar as observed for *P. pastoris* $\Delta ku70$, CRISPR/Cas9 expression (*CAS9* + gRNA) in the oleaginous yeast *Yarrowia lipolytica* $\Delta ku70$ and $\Delta ku70/\Delta ku80$ strains yielded reduced targeting and transformation efficiencies.⁴⁶ Based on these results, CRISPR/Cas9 vectors bearing ribozyme flanked gRNAs to target *GUT1* and *HsCAS9* were selected for the co-transformation experiments with homology donor cassettes,

because the highest targeting efficiencies/functionalities were obtained with these constructs in the $\Delta ku70$ strain.

CRISPR/Cas9 + donor DNA co-transformation in *P. pastoris* $\Delta ku70$ increases HDR

As next step *P. pastoris* $\Delta ku70$ and the wildtype strain were co-transformed with CRISPR/Cas9 plasmids and marker-free donor cassettes and selected for the presence of the episomal CRISPR/Cas9 plasmid: A *GUT1* donor cassette (previously reported³² consisting of two homologous arms directly fused together (total length 2 kbp, 1 μ g used for transformation)) and an episomal CRISPR/Cas9 plasmid (100 ng) to target the *GUT1* CDS were co-transformed (Figure 2A). Donor cassettes without a resistance marker allow to completely delete an ORF, to replace it with a different sequence or to insert sequences coding, for example, for a tag. The CRISPR/Cas9 plasmid can be simply removed by ARS plasmid loss by growing the strains on non-selective media.³² By applying Zeocin selection, only transformants, which contained a CRISPR/Cas9 plasmid, were able to grow on selective agar plates. Ideally the homologous recombination rate of the donor would be increased up to 100% by the site specific strand break, which would allow the use of marker-free homologous recombination cassettes for the replacement of genomic DNA. This turned out to be not the case in the wildtype strain, where NHEJ remained the preferred pathway for repair.³² However, high HR efficiencies for the marker-free donor cassette of 78% (GUT1-gRNA2) to 91% (GUT1-gRNA1) were obtained with the *P. pastoris* $\Delta ku70$ strain (Figure 2B). Note that also the *P. pastoris* $\Delta ku70$ strain alone, without Cas9 and gRNA permits integration rates of a control cassette containing a resistance marker of approximately 50% at the *GUT1* locus (Figure 2B). Using Cas9 and the gRNA to target *GUT1* further improved the homologous recombination rate 1.6- to 1.9-fold compared to the control (+donor DNA with Zeocin cassette, control without CRISPR/Cas9). Hence, the Cas9/gRNA system allowed to achieve marker less modifications even at a higher frequency than site specific marker integration in the $\Delta ku70$ strain.

We additionally used the control knockout cassette consisting of the Zeocin resistance cassette flanked by homology arms for CRISPR-Cas9 co-transformation experiments, in order to determine if the size of the donor DNA could have influenced the integration rates (2 kbp—donor DNA lacking the Zeocin marker cassette, 3.8 kbp—donor DNA bearing the Zeocin marker cassette). By applying Zeocin selection, it was not possible to discriminate transformants, which took up the CRISPR/Cas9 plasmid, the donor cassette or both recombinant DNA fragments. The HR efficiencies were in the same range as for the marker-free donor cassettes: 78% for GUT1-gRNA2 to 94% for GUT1-gRNA3 suggesting that this length difference does not have a clear influence. For both donor fragments the lowest HR efficiency was obtained with

GUT1-gRNA2, indicating that this gRNA was generally less suited for such genome targeting approach.

The transformation rates of *P. pastoris* $\Delta ku70$ using the donor fragments and CRISPR/Cas9 plasmids were increased (between 200 and 300 CFUs/ μ g CRISPR/Cas9 plasmid DNA independently from the donor cassette and gRNA used) compared to transformations, where only a CRISPR/Cas9 plasmid had been transformed (less than 60 CFUs/ μ g CRISPR/Cas9 plasmid DNA), but were as expected still considerably lower than the wildtype strain (5000 CFUs/ μ g CRISPR/Cas9 plasmid DNA). The presence of the homology repair fragment seemed to reduce the toxicity of the CRISPR/Cas9 system in the *P. pastoris* $\Delta ku70$ strain.

This system provides for the first time the possibility to achieve marker less modifications in *P. pastoris* at frequencies approaching 100%. In a typical experiment of transforming 100 ng of CAS9/gRNA plasmid + donor, we obtained a few dozen colonies and nearly all of them were positive. Also in the case of *Y. lipolytica*, where the HR frequencies of donor cassettes were low to moderate in the wildtype strain (8% for the *TRP1* locus, 73% for the *MFE1* locus), integration frequencies of 100% could be obtained with a $\Delta ku70$ strain.^{46,47}

3.1.2 | Cas9 nickase to increase HR frequencies in *P. pastoris* wildtype was not successful

Although HR efficiencies could be increased in *P. pastoris* $\Delta ku70$ and we were able to integrate marker-free donor cassettes, NHEJ remained the prevalent repair mechanism upon CRISPR/Cas9 targeted cleavage in *P. pastoris* wildtype strains. In order to increase HDR in *P. pastoris* wildtype strains, we constructed plasmids for the expression of the Cas9 nickase. This Cas9 variant creates single strand breaks due to the inactivation of one of its catalytic domains via the introduction of a point mutation.^{34,35} The Cas9 nickase, which was cloned into *P. pastoris* vectors, bore the amino acid exchange D10 \rightarrow A.¹⁵ By introducing single strand breaks, we intended to reduce the toxicity of the system and to mediate donor recombination, as demonstrated by Cong et al and Hsu et al.^{36,37} We generated plasmids containing *CAS9* nickase and either GUT1-gRNA0 or GUT1-gRNA4 to target the *GUT1* locus. These gRNAs direct the cleavage in close proximity of the donor recombination sites (gRNA0 bp -19-1 bp of *GUT1*, gRNA4 bp 1833-1853 bp [numbering relative to start codon as +1]), which would benefit the formation of single stranded homology arms (Figure 3A) Targeting efficiencies ($\Delta gut1$ phenotype) of 2% were obtained for GUT1-gRNA0 and of 64% for GUT1-gRNA4 in combination with the wildtype Cas9 nuclease. The transformation of CRISPR/Cas9 plasmids bearing Cas9 nickase and either GUT1-gRNA0 or GUT1-gRNA4 did not result in glycerol deficient *P. pastoris* cells and also the co-transformation of

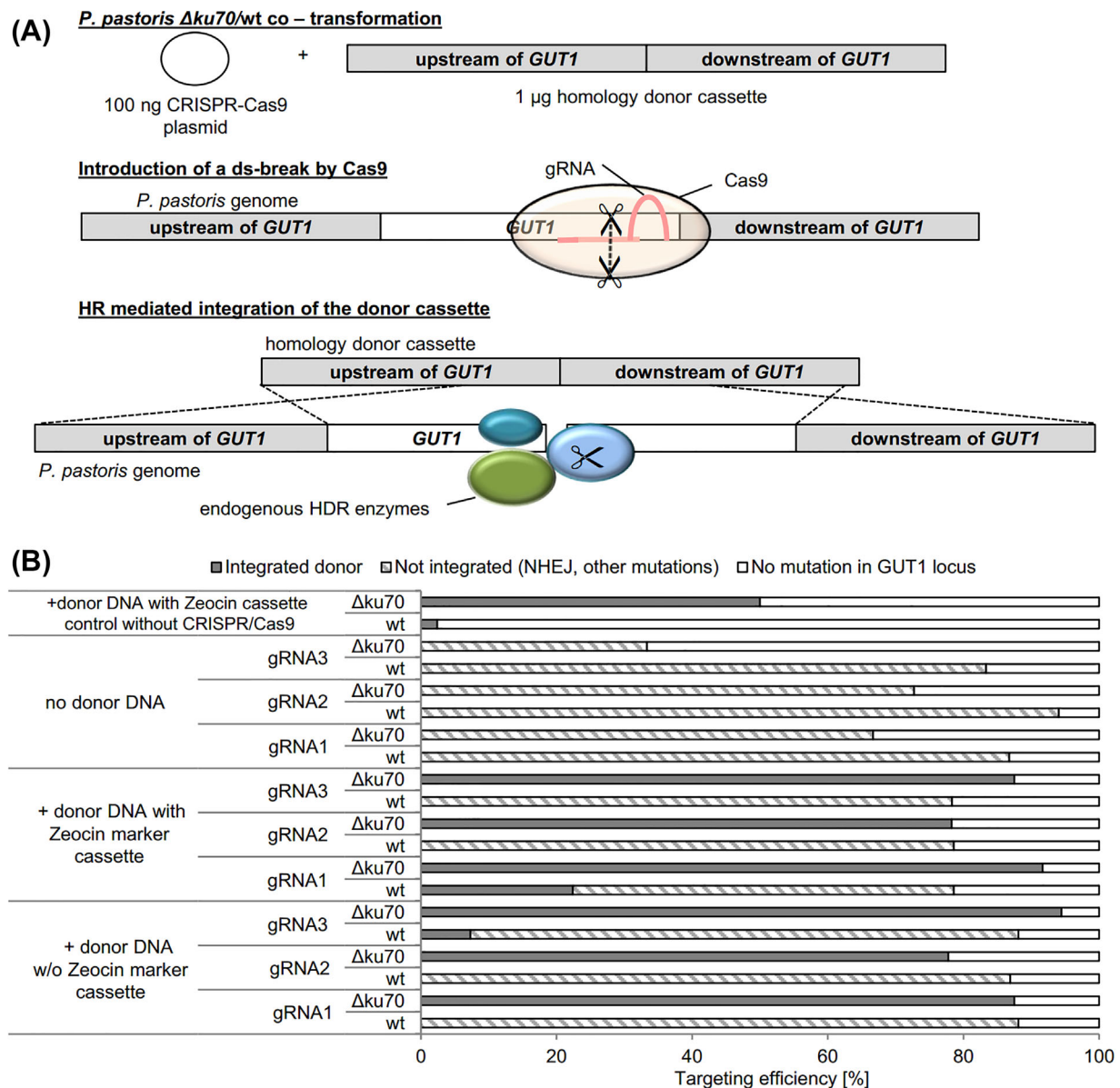


FIGURE 2 Markerless donor cassette integration with efficiencies approaching 100% in a *P. pastoris ku70* deletion strain. A, Principle of CRISPR/Cas9 mediated integration of donor cassettes to allow the complete deletion of genomic DNA. As first step *P. pastoris* $\Delta ku70$ was co-transformed with 100 ng CRISPR/Cas9 plasmid and a linear homologous donor cassette (1 μ g). The nuclease Cas9 introduced a double strand break at the target locus, which was the starting point for cellular repair. Host cell HR enzymes mediate the integration of the homologous donor fragment. CRISPR/Cas9 plasmids can be cured by growth on non-selective media. The elements are not drawn to scale. B, CRISPR/Cas9 mediated integration of homologous donor cassettes does not strongly affect the wildtype strain, but improves specific integration in the *P. pastoris* $\Delta ku70$ strain. Competent cells were co-transformed with 100 ng CRISPR/Cas9 circular plasmid DNA and 1 μ g donor DNA to target *GUT1*. *P. pastoris* CBS 7435 wildtype results have previously been reported.³² As control solely one μ g of the donor DNA was used for the transformation. Upon cultivating random transformants in 96-DWPs, cell material was transferred with a metallic stamp on minimal media agar plates with either glucose (BMD1) or glycerol (BMG1) as carbon source to determine the targeting efficiency. The integration of the donor cassettes was verified by sequencing of >10 glycerol deficient transformants per CRISPR/Cas9 plasmid—donor DNA combination. Mean values and standard deviations of biological triplicates are shown

HsCas9 nickase plasmids and donor cassettes rendered most transformants unaffected, indicating that the single strand breaks were repaired correctly and the donor cassette had not been used as repair template (Figure 3B). *P. pastoris* $\Delta ku70$

was not transformed with Cas9 nickase plasmids, because single strand breaks are repaired independently from NHEJ,^{26,48} and we did not expect differences compared to previous transformations.

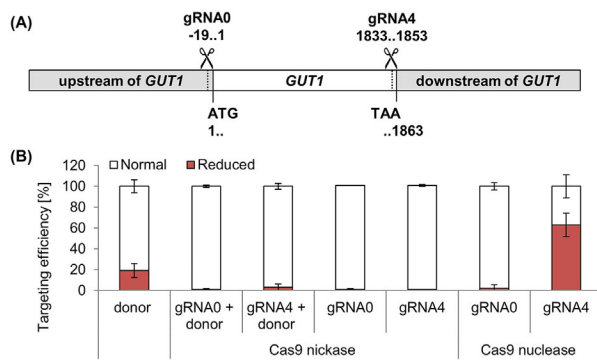


FIGURE 3 Use of the Cas9 nickase to increase HR frequencies in *P. pastoris* wildtype was not successful. A, *GUT1*-gRNA0 and *GUT1*-gRNA4 were designed, which should favor the formation of single stranded of homology arms for the integration of donor cassettes. B, *P. pastoris* CBS 7435 wildtype was transformed with CRISPR-Cas9 plasmids bearing either *HsCas9* or *HsCas9* nickase and a gRNA to target *GUT1* (*GUT1*-gRNA0 or *GUT1*-gRNA4). Transformants were cultivated in 96-DWPs and then stamped on minimal media plates containing glucose (BMD1) and glycerol (BMG1) as carbon source to determine the targeting efficiencies. The targeting efficiencies were only calculated by counting the glycerol deficient colonies and not by sequencing genomic DNA. Cas9 nickase transformants did not display a reduced growth phenotype compared to transformants with the Cas9 nuclease (*GUT1* deficient transformants obtained with *GUT1*-gRNA4). In case of *GUT1*-gRNA0/Cas9 nuclease either no cleavage had taken place due to a non-functional gRNA or the cleavage did not result in a visible phenotype. Co-transformations of the Cas9-nickase plasmids (100 ng) and homology donor fragments (1 μ g) yielded less *GUT1* deficient clones compared to transformations with solely the donor cassette

3.1.3 | Autonomous replication of donor fragments increases HR rates in *P. pastoris* wildtype strains

As final attempt to increase HR frequencies for the integration of marker less donor cassettes in *P. pastoris* wildtype strains, we constructed donor cassettes bearing autonomous replication sequences (ARSs, Figure 4). Similar as observed for *Yarrowia lipolytica*,⁴⁶ we assumed that the linear DNA used in the previous experiments (Figure 2) was not sufficiently sustained in the cells and that episomal maintenance might improve the HR frequency. Co-transformations of CRISPR/Cas9 plasmids (100 ng) and various donor fragments were performed. Either 100 ng circular donor DNA lacking an ARS, 100 ng circular donor DNA containing an ARS, 1000 ng circular donor DNA with ARS, 1000 ng *HindIII* linearized donor DNA with ARS, no donor DNA or 1000 ng PCR amplified donor DNA (without ARS) were used for the transformations (Figure 4A). When co-transforming a circular donor plasmid lacking an ARS, all *P. pastoris* wildtype transformants bore NHEJ mutations, similar to

the transformations, where the donor was omitted. Co-transforming 100 ng circular donor DNA containing an ARS yielded 92% clones with NHEJ mutations and a small proportion (6%), which integrated the donor fragment. When 1000 ng of the circular donor DNA containing an ARS were used, in 33% of the transformants the deletion of the *GUT1* CDS was observed by the integration of the homology donor. The HR frequencies could be further increased using the linearized donor DNA fragments containing the ARS (47% HR transformants, Figure 4B).

Interestingly, this experiment provided also insights on the recombination process in *P. pastoris*: The linear ends of the PCR amplified donor cassettes used as control are directly homologous to the genome (Figure 4A). In contrast, the ends of the linearized circular plasmid are different from the *GUT1* locus targeted, and the *GUT1* homologous sequences are placed relatively far (280 and 2917 bp) from the *HindIII* linearization site. Nonetheless, 47% of the constructs were integrated correctly. These results suggest, that the HR mechanism in *P. pastoris* “searches” for complementary sequences to the DSB introduced by Cas9/gRNA in the genome. Even more interestingly, also the circular plasmid yielded 33% correct integration. In this case, no linear ends are provided on the plasmid and the recombination event must be initiated by the DSB in the genome. Transformation efficiencies of ARS containing donors appeared to be slightly reduced. While the co-transformation of 100 ng circular donor without ARS yielded \sim 5000 cfu/ μ g CRISPR/Cas9 plasmid DNA, the transformation efficiency decreased to \sim 1000 cfu/ μ g CRISPR/Cas9 plasmid DNA with 100 ng of circular donor containing an ARS. In case of the linearized ARS containing donor, the transformation efficiency decreased slightly more to \sim 500 cfu/ μ g CRISPR/Cas9 plasmid DNA. Larger amounts of DNA in the transformation (and depending on the form circular/linear) hence appear to influence transformation rates in *P. pastoris*. Yet, the applicability of ARS containing donor fragments for genome engineering is not impaired, as the transformation rates are still sufficient to obtain many correct transformants (as discussed similarly for the $\Delta ku70$ strain without ARS cassettes).

To our knowledge, this is the first work showing the integration of marker-less donor cassettes at reasonable rates in *P. pastoris* wildtype strains. These kinds of ARS donor cassettes are especially valuable for the replacement or removal of DNA sequences, where artifacts such as scars generated by site specific recombinases or selection marker cassettes might influence their functionality (eg, regulatory elements in promoters or “tagged” coding sequences). ARSs might be also placed on classical knockout cassettes to guarantee the maintenance of the cassette throughout the cell cycle, especially in the S/G2 phase where HR is predominant.⁴⁹

When co-transforming the NHEJ deficient $\Delta ku70$ strain with the CRISPR/Cas9 plasmid and 100 ng circular donor

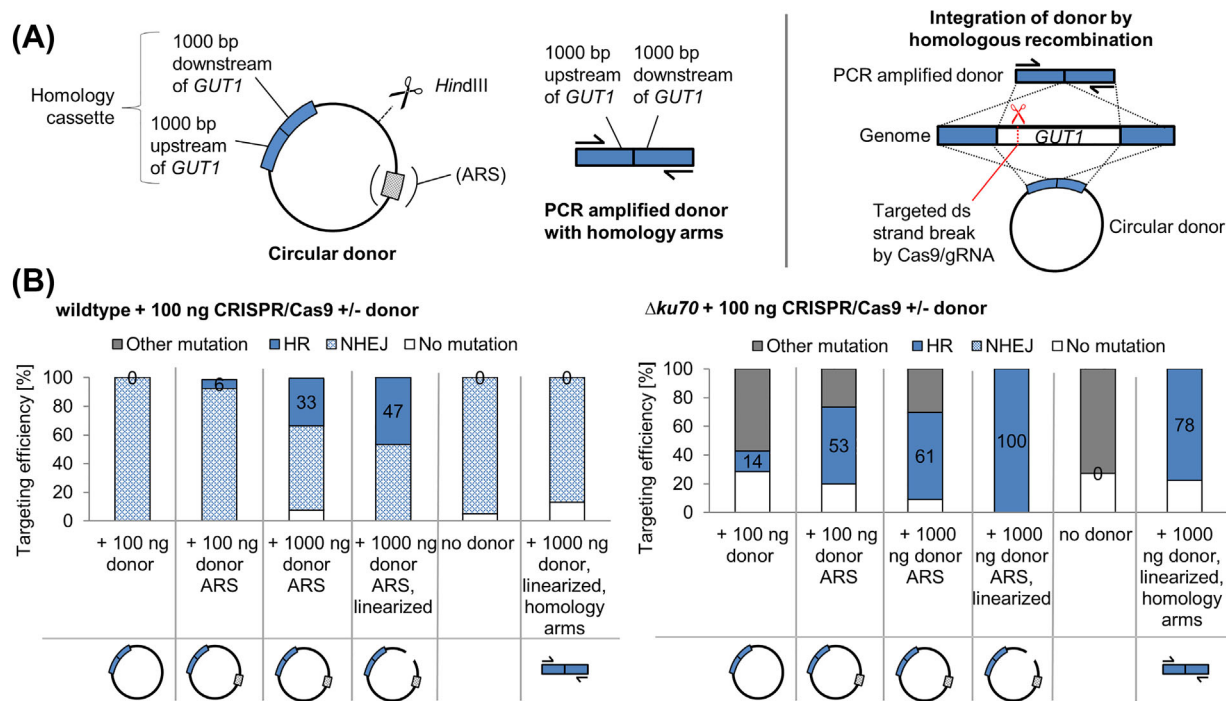


FIGURE 4 ARSs promote efficient CRISPR/Cas9 mediated integration of markerless donor fragments in the *P. pastoris* wildtype strain. A, Design of the donor cassettes (left side) and recombination events targeted (right side). A circular plasmid was used for the *P. pastoris* transformation, which contains 1000 bp regions 5' upstream and 3' downstream of *GUT1* (and an ARS for replication in yeast). As control the PCR amplified homology cassette consisting of the 1000 bp regions 5' upstream and 3' downstream of *GUT1* was used for the transformation. Due to sequence homology of the donor cassette the *GUT1* CDS can be removed by homologous recombination. B, *P. pastoris* wildtype (left side) and $\Delta ku70$ (right side) co-transformation with 100 ng CRISPR-Cas9 plasmid DNA (*GUT1*-gRNA2) and various donor fragments: 100 ng circular donor DNA lacking ARS, 100 ng circular donor DNA with ARS, 1000 ng circular donor DNA with ARS, 1000 ng circular donor DNA with ARS linearized with *HindIII*, no donor DNA and 1000 ng PCR amplified donor DNA. Results on the PCR amplified donor cassette in the wildtype strain have previously been reported.³² Upon cultivating random transformants in 96-DWPs, cell material was transferred with a metallic stamp on minimal media agar plates with either glucose (BMD1) or glycerol (BMG1) as carbon source to determine the targeting efficiency. The percentages of HR integrated donor fragments are denoted as numbers in the bar diagram. The integration of the donor cassettes was verified by sequencing of >10 glycerol deficient transformants per CRISPR/Cas9 plasmid—donor DNA combination. The ARS-donor enables the integration of marker-less donor fragments in almost 50% of the transformants, when 1000 μ g linear donor DNA were used

DNA lacking an ARS, 14% of the transformants integrated the donor fragment, which resulted in the deletion of the *GUT1* CDS (Figure 4B). The co-transformation of *P. pastoris* $\Delta ku70$ with the CRISPR/Cas9 plasmid and 100 ng circular donor DNA containing an ARS yielded 53% HR transformants, 1000 ng circular donor DNA with ARS 61% HR transformants, and 1000 ng linearized donor DNA with ARS rendered all transformants glycerol deficient. However, in contrast to the *P. pastoris* wildtype strain, the PCR amplified donor bearing homology arms was also integrated at a high rate (78%), similar to transformations with the ARS donor cassettes. Also transformation rates with ARS containing donor fragments remained similarly low in the $\Delta ku70$ strain (50-100 CFUs/ μ g CRISPR/Cas9 plasmid DNA) as when transforming regular PCR products (section 3.1.1). Hence while providing a major advantage in the wildtype strain,

using ARS on the donor did not lead to clear improvements in the $\Delta ku70$ strain.

3.1.4 | Reduction of toxicity of the CRISPR/Cas9 system in *P. pastoris* using hf-Cas9 was not successful

In order to increase the transformation efficiency and reduce the toxicity of the CRISPR/Cas9 system, we tried to implement a high fidelity Cas9 variant in *P. pastoris*. High fidelity Cas9 nucleases were generated by structure-guided protein engineering and reduce off-target effects, while maintaining robust on-target cleavage.^{18,19} Wildtype Cas9 can cleave off-target sites that are not fully complementary to the gRNA, when the binding strength to the non-complementary DNA strand is high. If interactions between

Cas9 and the non-complementary DNA strand are reduced due to amino acid exchanges, an increase in the specificity had been observed. Four amino acids in the Cas9 sequence were exchanged according to Kleinstiver et al¹⁸ to prepare the off-target free version of Cas9 (497N → A, 661R → A, 695Q → A, 926Q → A), hf-Cas9. The hf-Cas9 coding sequence was cloned into the vector previously used (bearing GUT1-gRNA1 flanked by ribozymes under the control of the bidirectional *HTX1* promoter). The targeting efficiency for the *GUT1* locus was similar (80%), compared to the transformations using unmutated, wildtype Cas9 (91%), indicating that on-target activities had been retained (data not shown). The indel mutations in the *GUT1* locus introduced by either Cas9 or hf-Cas9 were verified by sequencing. In two clones bearing Cas9, a single bp (G-C) located four bps upstream of the protospacer adjacent motif (PAM, for Cas9: AGG) was deleted (CGAGTACTC-TACCTCTCTCAGG, 2/2 sequenced), as well as in one of the sequenced clones expressing hf-*HsCAS9* (1/2 sequenced). In the second hf-*HsCAS9* transformant, one bp located 3 bp upstream PAM was deleted (CGAGTACTCTACCTCTGT-CAGG). The PAM is a 3 bp DNA sequence immediately following the Cas9 nuclease target DNA sequence. PAM is essential for Cas9 mediated DNA unwinding and subsequent DNA cleavage.⁵⁰ These initial implementation studies were performed in a *P. pastoris* wildtype strain. As next step we transformed *P. pastoris* $\Delta ku70$ where transformation efficiencies were 10-fold reduced due to the impaired NHEJ-repair pathway (Figure 1). Possible off-target cleavage is an additional burden to the CRISPR/Cas9 transformants. Thus we expected an increase in the transformation efficiency, if solely on-target cleavage had taken place (and was a major factor affecting the CRISPR/Cas9 efficiency in *P. pastoris*). The transformations of *P. pastoris* wildtype cells with either hf-Cas9 or Cas9 and GUT1-gRNA1 yielded 5000 CFUs/ μ g DNA. Similarly, when transforming *P. pastoris* $\Delta ku70$, transformation efficiencies of 20-40 CFUs/ μ g DNA were obtained with either hf-Cas9 or Cas9 and GUT1-gRNA1. Additional co-transformations with homology donor cassettes yielded similar transformation efficiencies with Cas9 and hf-Cas9 (20-60 CFUs/ μ g CRISPR-Cas9 plasmid DNA). Thus inferred from this data, hf-Cas9, generated in this work, appears to show similar on- and off-target activities compared to the parental Cas9 enzyme in our experimental setup in *P. pastoris*. The use of hf-Cas9 does not appear to provide advantages over the wildtype Cas9 enzyme in *P. pastoris*.

3.2 | DNA deletion by multiplexing

Besides gene replacement by the co-transformation of donor fragments, DNA removal was investigated using gRNAs, which direct multiple Cas9 cleavages in close proximity in the *P. pastoris* genome. Thereby sequences from the genome

could be removed without requiring a donor cassette to repair the gap (Figure 5). The *DAS1/DAS2* locus was selected for the implementation. In *P. pastoris* two isoforms of the dihydroxyacetone synthase, *DAS1* and *DAS2*, occur, which are 91% identical.⁵¹⁻⁵³ These genes are located in the reverse complement orientation and are interspaced with the respective promoters and *HOB3*, encoding a putative guanosine nucleotide exchange factor. *P. pastoris* *DAS1* and *DAS2* double knockout strains display a reduced growth on methanol similar to *AOX1* deficient mut^S strains. Whereas single knockouts have a similar growth rate on methanol as sole carbon source, compared to *P. pastoris* wildtype strains.⁵³ Due to the high sequence identity a single gRNA can be designed, which directs cleavage in both, the *DAS1* and the *DAS2* gene (Figure 5).

Two different gRNAs (*DAS1/DAS2*-gRNA1 and *DAS1/DAS2*-gRNA2) were cloned in CRISPR/Cas9 plasmids bearing *HsCas9*. The *DAS1/DAS2*-gRNA1 cleavage sites were interspaced by a 3224 bp fragment, whereas the *DAS1/DAS2*-gRNA2 cleavage sites were interspaced by a 6293 bp fragment (Figure 5A). The overall targeting efficiencies for the simultaneous cleavage at the *DAS1* and *DAS2* locus were similar for both gRNAs (87% for *DAS1/DAS2*-gRNA1 and 86% for *DAS1/DAS2*-gRNA2). However the type of mutation differed depending on the distance of cleavage sites (Figure 5B). Shorter distances led to the removal of DNA fragments, however, chromosomal rearrangements and translocations were observed rather than clean cuts. PCR amplifying and sequencing of these transformants resulted either in no PCR product or sequencing reads, which could only be aligned partially. Local translocations might be favored due to the palindromic nature of the target sequences. *DAS1/DAS2*-gRNA1 mediated cleavage appeared to predominantly trigger rearrangement/translocation events at the *DAS1/DAS2* locus (79%). When using *DAS1/DAS2*-gRNA2 48% of the transformants bore rearrangement/translocation mutations, and the number of transformants containing point mutations at the Cas9 cleavage site increased to 38.6% compared to *DAS1/DAS2*-gRNA1 transformants (8.8%).

3.3 | CRISPR/Cas9 expression from geneticin marker plasmids

P. pastoris production strains often contain a Zeocin cassette integrated in the genome, which was required for the initial selection for the gene of interest. At later protein production stages the Zeocin resistance is not required anymore, since the genomic integration of expression cassettes results in stable expression strains, which maintain the integrated cassette in the absence of selective pressure. However re-transformations of these strains with plasmids containing the Zeocin resistance marker, for example, for the expression of

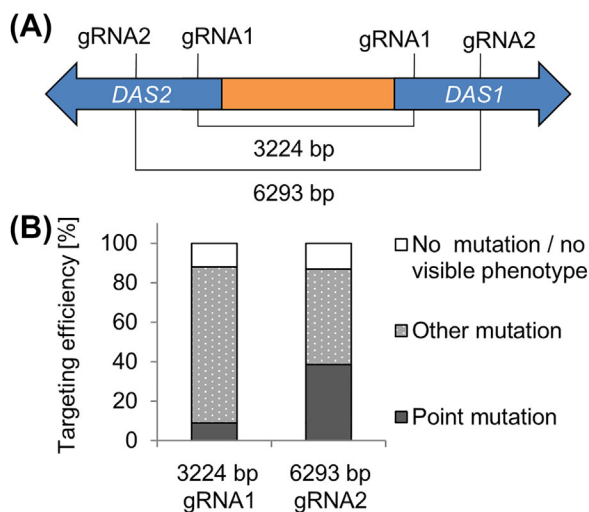


FIGURE 5 Dual cutting gRNAs for DNA removal. A, The *P. pastoris* *DAS1/DAS2* locus consisting of *DAS1* and *DAS2*, which are located in the reverse complement orientation and are interspaced with the respective promoters and *HOB3*, encoding a putative guanosine nucleotide exchange factor. Due to 91% sequence identity a single gRNA can be designed to introduce double strand breaks in the *DAS1* and *DAS2* gene. B, *P. pastoris* CBS 7435 wildtype was transformed with CRISPR-Cas9 plasmids bearing *HsCas9* and either *DAS1/DAS2*-gRNA1 or *DAS1/DAS2*-gRNA2. Transformants were cultivated in 96-DWPs and then stamped on minimal media plates containing glucose (BMD1) and methanol (BMM1) as carbon source to determine the targeting efficiencies. The type of mutation was elucidated by sequencing of >10 methanol deficient transformants/construct. Mean values and standard deviations of biological triplicates are shown. Increasing the distance of the targeting sites increases the frequencies of NHEJ mutations and reduces the probability for sequence removal

helper genes/chaperones and also CRISPR/Cas9 mediated strain engineering with the plasmids currently established³² cannot be performed.

In order to expand the CRISPR/Cas9 toolset for *P. pastoris* we generated a CRISPR/Cas9-Geneticin plasmid bearing a Geneticin resistance cassette, which can be used to introduce targeted mutations in preexisting Zeocin resistant *P. pastoris* strains. Moreover we also applied this plasmid for marker recycling⁵⁴ by transforming Zeocin resistant *P. pastoris* strains with CRISPR/Cas9-Geneticin plasmids containing gRNAs to target the Zeocin resistance gene (Figure 6). Using this approach Zeocin sensitive *P. pastoris* strains were generated, which can be re-transformed with plasmids bearing Zeocin resistance markers.

As first step, the CRISPR/Cas9-Geneticin system was compared to the already published CRISPR/Cas9-Zeocin vectors by targeting the *GUT1* locus (Figure 6A). The *CAS9*/gRNA expression cassette of the newly generated CRISPR/Cas9-Geneticin plasmid were identical to the

Zeocin based designs used previously³² and in this study. Only the resistance marker cassette was exchanged. We transformed a *P. pastoris* wildtype strain with CRISPR/Cas9 plasmids lacking the gRNAs, and harboring either the Zeocin resistance marker or the Geneticin resistance marker. Using the CRISPR/Cas9-Zeocin and the CRISPR/Cas9-Geneticin plasmid more than 20 000 CFUs/ μ g DNA were obtained. However, it may be noted that in previous experiments a reduced number of CRISPR/Cas9-Geneticin transformants had been observed (data not shown). Transformation efficiencies of 3280 to 4735 CFUs/ μ g DNA were obtained with the CRISPR/Cas9-Geneticin plasmids containing gRNAs. Using these plasmids targeting efficiencies of 85-100% for the *GUT1* locus similar to the transformations of the CRISPR/Cas9-Zeocin plasmids (87-94% for *GUT1* and different gRNAs,³²) were obtained (Figure 6A). Thus the described CRISPR/Cas9 experiments with *P. pastoris* can be performed with Gen/KanMX based vectors as well as with vectors using the Zeocin resistance gene.

For the marker recycling and the for inactivating the resistance marker gene in existing Zeocin based *P. pastoris* strains, different gRNAs were designed (ZEO-gRNA1, ZEO-gRNA2, ZEO-gRNA3) to introduce indels in the CDS of the Zeocin resistance gene and to inactivate the protein functionality (Figure 6B). *P. pastoris* strains harboring plasmids for the expression of either enhanced green fluorescent protein (eGFP), *C. antarctica* lipase B (CaLB) or glucose oxidase (Gox) with Zeocin resistance markers were transformed with CRISPR/Cas9-Geneticin plasmids to target the CDS of the Zeocin resistance gene. The targeting efficiencies varied depending on the gRNA used: ZEO-gRNA1 yielded the lowest targeting efficiencies in the three *P. pastoris* strains (0-11%). Moderate targeting efficiencies were obtained with ZEO-gRNA2 (20-50%), whereas ZEO-gRNA3 enabled highly efficient gene targeting (62-100%) (Figure 6B). CRISPR/Cas9 plasmid curing was performed by growing the strains on non-selective media. This strategy can be adapted for the recycling of various dominant selection markers including antibiotics and auxotrophy markers, and offers an alternative to the existing recombinase systems (Flippase/FRT, Cre/loxP, reviewed by Vogl et al⁵⁴), where the recognition sites have to be incorporated in the plasmid prior to the transformation.¹² In addition, also two gRNAs could be used to completely excise the marker gene, as discussed for the *DAS1/2* locus (section 3.2 and Figure 3). However, note that the coding sequence of the Zeocin resistance gene of the pPpT4 plasmids family,²³ that was used as the basis in our work, is different from other commonly used vectors (eg, Invitrogen vectors). Hence the gRNAs reported here for Zeocin marker recycling would possibly need to be adapted for the use with other vector families.

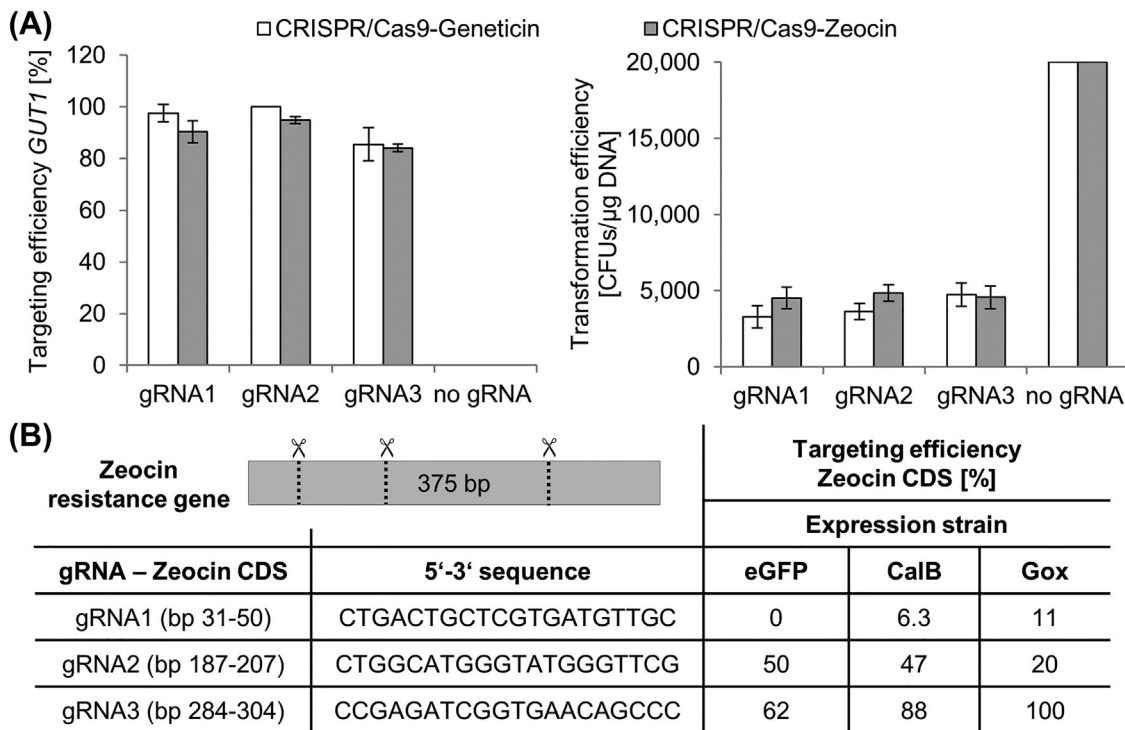


FIGURE 6 Characterization of the CRISPR/Cas9-Geneticin system for *P. pastoris* and application for marker recycling of Zeocin resistance cassette containing strains. A, *P. pastoris* cells were transformed with CRISPR/Cas9 plasmids containing gRNAs to target *GUT1* and a Geneticin resistance marker cassette. The transformants were cultivated in 96-DWPs and stamped on minimal media plates with different carbon sources (glucose [BMD1] and glycerol [BMG1]) to determine the targeting efficiency. The Geneticin results are compared to the CRISPR/Cas9-Zeocin system published previously.³² Mean values and standard deviation of biological triplicates are shown. In case of the Geneticin *GUT1*-gRNA2, all replicates yielded 100% correct integration, hence no standard deviation is shown. B, The CRISPR/Cas9-Geneticin plasmid was used for marker recycling in existing *P. pastoris* production strains. Several other *P. pastoris* strains bearing a Zeocin resistance cassette and a reporter gene/enzyme expression cassette (eGFP, CalB, Gox) integrated in the genome were transformed with CRISPR/Cas9 plasmids harboring gRNAs to target the Zeocin resistance gene of the pPpT4 vector.²³ Indel mutations were introduced in the CDS of the Zeocin resistance gene, which rendered the strains Zeocin sensitive. The targeting efficiencies were determined by stamping the randomly picked transformants on YPD and YPD-Zeocin (50 μ g/mL) agar plates. Mean values and standard deviations of biological triplicates are shown

4 | CONCLUSION AND OUTLOOK

Continuous developments of CRISPR/Cas9 systems broaden the range of applications to engineer expression hosts including the methylotrophic yeast *P. pastoris*. In this work, efficient and precise CRISPR/Cas9 induced HR was on the one hand demonstrated by transforming the NHEJ-impeded *P. pastoris* $\Delta ku70$ strain and on the other hand also efficiencies for the *P. pastoris* wildtype strain were improved by using donor cassettes with ARSs. Thereby also homology fragments lacking selection marker cassettes were correctly integrated at high frequencies. In case of the *P. pastoris* $\Delta ku70$ strain low transformation rates were obtained. However, in a typical experiment of transforming 100 ng CRISPR/Cas9 plasmid DNA and 1 μ g linear donor DNA resulted in a few dozen colonies and nearly all of them were positive. Inhibiting

NHEJ to increase the HR efficiency in combination with CRISPR/Cas9 gene targeting had been successfully applied in alternative yeast species,^{28,46,47} mammalian cell lines, and mice⁵⁵ before and is now for the first time applicable for *P. pastoris* $\Delta ku70$ strains. Furthermore, for the integration of homology donor fragments in a *P. pastoris* wildtype strain the addition of ARSs increased the integration rates from as little as 0% up to approximately 50%. To our knowledge this is the first report demonstrating the integration of markerless donor cassettes at reasonable rates in *P. pastoris* wildtype strain, which enables novel engineering strategies. Markerless donor cassettes are especially valuable for the replacement or removal of DNA sequences, where artifacts such as scars generated by site specific recombinases or selection marker cassettes might influence their functionality (eg, regulatory elements in promoters or “tagged” coding

sequences). ARSs might be also placed on classical knockout cassettes to guarantee the maintenance of the cassette throughout the cell cycle, especially in the S/G2 phase where HR is predominant and thereby favor their integration.⁴⁹ Hence with these systems, either close to 100% correct integration can be achieved using the $\Delta ku70$ strain (at the expense of total transformation efficiency) or the wildtype strain can be used allowing ~50% correct integration. To expand the CRISPR/Cas9 toolset of *P. pastoris* we also generated CRISPR-Cas9 plasmids, where transformants can be selected with Geneticin. Those can also be applied to introduce targeted mutations in Zeocin resistant *P. pastoris* strains. Moreover CRISPR/Cas9 plasmids with different selection markers enable marker recycling by using alternating gRNAs to delete resistance marker genes. The results presented here may aid to make the CRISPR/Cas9 technology even more appealing to the *P. pastoris* community, provide a deepened understanding of CRISPR/Cas9 for strain engineering and might be adapted for alternative production hosts. For the first time also to the introduction of scarless tags, point mutations in genome, protein fusions or deletions of whole genome sequence stretches will be enabled employing these tools.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the Innovative Medicines Initiative Joint Undertaking project CHEM21 under grant agreement n°115360, resources of which are composed of financial contribution from the European Unions “Seventh Framework Programme (FP7/2007-2013) and EFPIA companies” in kind contribution. Moreover the research for this work has received funding from the European Union (EU) project ROBOX (grant agreement n° 635734) under EU’s Horizon 2020 Programme Research and Innovation actions H2020-LEIT BIO-2014-1. The views and opinions expressed are only those of the author(s), and do not necessarily reflect those of the European Union Research Agency. The European Union is not liable for any use that may be made of the information contained here. T.V. gratefully acknowledges the Austrian Science Fund (FWF) project number W901 (DK “Molecular Enzymology” Graz) and we acknowledge support from NAWI Graz.

AUTHORS’ CONTRIBUTIONS

AW and JF contributed equally. AW, JF, and TV designed the experiments. AW, JF, HR, and CK performed the experiments. AW, JF, and TV interpreted the data and wrote the manuscript. AG, AW, JF, and TV conceived of the study.

AG and TV supervised the research. All authors read and approved the final manuscript.

ORCID

Thomas Vogl  <http://orcid.org/0000-0002-3892-1740>

REFERENCES

1. Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotechnol.* 2014;32:347–355.
2. Wang H, La Russa M, Qi LS. CRISPR/Cas9 in genome editing and beyond. *Ann Rev Biochem.* 2016;85:227–264.
3. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science.* 2014;346:1258096.
4. Xue W, Chen S, Yin H, et al. CRISPR-mediated direct mutation of cancer genes in the mouse liver. *Nature.* 2014;514:380–384.
5. Matano M, Date S, Shimokawa M, et al. Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. *Nat Med.* 2015;21:256–262.
6. Wu Y, Liang D, Wang Y, et al. Correction of a genetic disease in mouse via use of CRISPR-Cas9. *Cell Stem Cell.* 2013;13:659–662.
7. Baltimore D, Berg P, Botchan M, et al. Biotechnology. A prudent path forward for genomic engineering and germline gene modification. *Science* 2015;348:36–38.
8. Rooiintan A, Morowvat MH. Road to the future of systems biotechnology: CRISPR-Cas-mediated metabolic engineering for recombinant protein production. *Biotechnol Genet Eng Rev.* 2016;32:74–91.
9. Krappmann S. CRISPR-Cas9, the new kid on the block of fungal molecular biology. *Med Mycol.* 2017;55:16–23.
10. Shen C-C, Sung L-Y, Lin S-Y, Lin M-W, Hu Y-C. Enhancing protein production yield from chinese hamster ovary cells by CRISPR interference. *ACS Synth Biol.* 2017;6:1509–1519.
11. Park G, Ouyang S, Borkovich KA. High-throughput construction of genetically modified fungi. In: Schmoll M, Dattenböck C, editors. *Gene Expression systems in fungi: advancements and applications.* Cham: Springer International Publishing; 2016: pp 241–252.
12. Weninger A, Killinger M, Vogl T. Key methods for synthetic biology: genome engineering and DNA assembly. In: Glieder A, Kubicek CP, Mattanovich D, Wiltschi B, Sauer M, editors. *Synthetic biology.* Vol 8. Cham: Springer International Publishing; 2016:pp 101–141.
13. Gaj T, Gersbach CA, Barbas CF. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* 2013;31:397–405.
14. Kim H, Kim J-S. A guide to genome engineering with programmable nucleases. *Nat Rev Genet.* 2014;15:321–334.
15. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science.* 2012;337:816–821.
16. Ma D, Liu F. Genome editing and its applications in model organisms. *Genomics Proteomics Bioinformatics.* 2015;13:336–344.

17. Kumar V, Jain M. The CRISPR-Cas system for plant genome editing: advances and opportunities. *J Exp Bot.* 2015;66:47–57.
18. Kleinstiver BP, Pattanayak V, Prew MS, et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature.* 2016;529:490.
19. Slaymaker IM, Gao LY, Zetsche B, Scott DA, Yan WX, Zhang F. Rationally engineered Cas9 nucleases with improved specificity. *Science.* 2016;351:84–88.
20. Bill RM. Playing catch-up with *Escherichia coli*: using yeast to increase success rates in recombinant protein production experiments. *Front Microbiol.* 2014;5:85.
21. Cregg JM, Cereghino JL, Shi J, Higgins DR. Recombinant protein expression in *Pichia pastoris*. *Mol Biotechnol.* 2000;16:23–52.
22. Ahmad M, Hirz M, Pichler H, Schwab H. Protein expression in *Pichia pastoris*: recent achievements and perspectives for heterologous protein production. *Appl Microbiol Biotechnol.* 2014;98:5301–5317.
23. Näätäsaari L, Mistlberger B, Ruth C, Hajek T, Hartner FS, Glieder A. Deletion of the *Pichia pastoris* KU70 homologue facilitates platform strain generation for gene expression and synthetic biology. *PLoS ONE.* 2012;7:e39720.
24. Smith F, Rouet P, Romanienko PJ, Jasin M. Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. *Nucleic Acids Res.* 1995;23:5012–5019.
25. Storici F, Durham CL, Gordenin DA, Resnick MA. Chromosomal site-specific double-strand breaks are efficiently targeted for repair by oligonucleotides in yeast. *Proc Natl Acad Sci USA.* 2003;100:14994–14999.
26. Caldecott KW. Single-strand break repair and genetic disease. *Nat Rev Genet.* 2008;9:619–631.
27. Jacobs JZ, Ciccaglione KM, Tournier V, Zaratiegui M. Implementation of the CRISPR-Cas9 system in fission yeast. *Nat Commun.* 2014;5:1–5.
28. Horwitz AA, Walter JM, Schubert MG, et al. Efficient multiplexed integration of synergistic alleles and metabolic pathways in yeasts via CRISPR-Cas. *Cell Syst.* 2015;1:88–96.
29. Vyas VK, Barrasa MI, Fink GR. A *Candida albicans* CRISPR system permits genetic engineering of essential genes and gene families. *Sci Adv.* 2015;1:e1500248.
30. Min K, Ichikawa Y, Woolford CA, Mitchell AP. *Candida albicans* gene deletion with a transient CRISPR-Cas9 system. *mSphere.* 2015;1:1–9.
31. Enkler L, Richer D, Marchand AL, Ferrandon D, Jossinet F. Genome engineering in the yeast pathogen *Candida glabrata* using the CRISPR-Cas9 system. *Sci Rep.* 2016;6:35766.
32. Weninger A, Hatzl A-M, Schmid C, Vogl T, Glieder A. Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast *Pichia pastoris*. *J Biotechnol.* 2016;235: 139–149.
33. Cregg JM, Barringer KJ, Hessler AY, Madden KR. *Pichia pastoris* as a host system for transformations. *Mol Cell Biol.* 1985;5: 3376–3385.
34. Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell.* 2014;157:1262–1278.
35. Ran FA, Hsu PD, Lin C, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell.* 2013;154:1380–1389.
36. Cong L, Ran FA, Cox D, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science.* 2013;339:819–823.
37. Hsu PD, Scott DA, Weinstein JA, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol.* 2013;31:827–832.
38. Sturmberger L, Chappell T, Geier M, et al. Refined *Pichia pastoris* reference genome sequence. *J Biotechnol.* 2016;235:121–131.
39. Krainer FW, Gmeiner C, Neutsch L, et al. Knockout of an endogenous mannosyltransferase increases the homogeneity of glycoproteins produced in *Pichia pastoris*. *Sci Rep.* 2013;3:3279.
40. Gibson DG, Young L, Chuang R, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods.* 2009;6:343–345.
41. Weninger A, Glieder A, Vogl T. A toolbox of endogenous and heterologous nuclear localization sequences for the methylotrophic yeast *Pichia pastoris*. H. Ah Kang, ed. *FEMS Yeast Res.* 2015;15: p.fov082.
42. Vogl T, Sturmberger L, Kickenweiz T, et al. A toolbox of diverse promoters related to methanol utilization: functionally verified parts for heterologous pathway expression in *Pichia pastoris*. *ACS Synth Biol.* 2016;5:172–186.
43. Lieber MR. NHEJ and its backup pathways in chromosomal translocations. *Nat Struct Mol Biol.* 2010;17:393–395.
44. Bétermier M, Bertrand P, Lopez BS. Is non-homologous end-joining really an inherently error-prone process? *PLoS Genetics* 2014;10:e1004086.
45. Brunet E, Simsek D, Tomishima M, et al. Chromosomal translocations induced at specified loci in human stem cells. *Proc Natl Acad Sci USA.* 2009;106:10620–10625.
46. Gao S, Tong Y, Wen Z, et al. Multiplex gene editing of the *Yarrowia lipolytica* genome using the CRISPR-Cas9 system. *J Ind Microbiol Biotechnol.* 2016;43:1085–1093.
47. Schwartz CM, Shabbir-Hussain M, Blenner M, Wheeldon I. Synthetic RNA polymerase III promoters facilitate high-efficiency CRISPR-Cas9-mediated genome editing in *Yarrowia lipolytica*. *ACS Synth Biol.* 2016;5:356–359.
48. Saffran WA, Greenberg RB, Thaler-scheer MS, Jones MM. Single strand and double strand DNA damage-induced reciprocal recombination in yeast. Dependence on nucleotide excision repair and RAD1 recombination. *Nucleic Acids Res.* 1994;22: 2823–2829.
49. Tsakraklides V, Brevnova E, Stephanopoulos G, Shaw AJ. Improved gene targeting through cell cycle synchronization. *PLoS ONE.* 2015;10:1–11.
50. Anders C, Niewoehner O, Duerst A, Jinek M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature.* 2014;513:569–573.
51. Krainer FW, Dietzsch C, Hajek T, Herwig C, Spadiut O, Glieder A. Recombinant protein expression in *Pichia pastoris* strains with an engineered methanol utilization pathway. *Microb Cell Fact.* 2012;11:22.
52. Vogl T, Glieder A. Regulation of *Pichia pastoris* promoters and its consequences for protein production. *New Biotechnol.* 2013;30: 385–404.
53. Geier M, Brandner C, Strohmeier GA, Hall M, Hartner FS, Glieder A. Engineering *Pichia pastoris* for improved NADH regeneration: a novel chassis strain for whole-cell catalysis. *Beilstein J Org Chem.* 2015;11:1741–1748.
54. Vogl T, Hartner FS, Glieder A. New opportunities by synthetic biology for biopharmaceutical production in *Pichia pastoris*. *Curr Opin Biotechnol.* 2013;24:1094–1101.

55. Maruyama T, Dougan SK, Truttmann MC, Bilate AM, Ingram JR, Ploegh HL. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotech.* 2015;33:538–542.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Weninger A, Fischer JE, Raschmanová H, Kniely C, Vogl T, Glieder A. Expanding the CRISPR/Cas9 toolkit for *Pichia pastoris* with efficient donor integration and alternative resistance markers. *J Cell Biochem.* 2018; 119:3183–3198. <https://doi.org/10.1002/jcb.26474>