CasPER, a method for directed evolution in genomic contexts using mutagenesis and CRISPR/Cas9

Jakociunas, Tadas; Pedersen, Lasse Ebdrup; Lis, Alicia Viktoria; Jensen, Michael Krogh; Keasling, Jay

Published in:
Metabolic Engineering

Link to article, DOI:
10.1016/j.ymben.2018.07.001

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Metabolic Engineering 48 (2018) 288–296

Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/meteng

CasPER, a method for directed evolution in genomic contexts using mutagenesis and CRISPR/Cas9

Tadas Jakočiūnas[a], Lasse E. Pedersen[a], Alicia V. Lis[a], Michael K. Jensen[a,*], Jay D. Keasling[a,b,c,d,e]

[a] The Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Denmark
[b] Joint BioEnergy Institute, Emeryville, CA, USA
[c] Biological Systems & Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[d] Department of Chemical and Biomolecular Engineering & Department of Bioengineering University of California, Berkeley, CA, USA
[e] Center for Synthetic Biochemistry, Institute for Synthetic Biology, Shenzhen Institutes for Advanced Technologies, Shenzhen, China

A B S T R A C T

Here we describe a method for robust directed evolution using mutagenesis of large sequence spaces in their genomic contexts. The method employs error-prone PCR and Cas9-mediated genome integration of mutant libraries of large-sized donor variants into single or multiple genomic sites with efficiencies reaching 98–99%. From sequencing of genome integrants, we determined that the mutation frequency along the donor fragments is maintained evenly and successfully integrated into the genomic target loci, indicating that there is no bias of mutational load towards the proximity of the double strand break. To validate the applicability of the method for directed evolution of metabolic gene products we engineered two essential enzymes in the mevalonate pathway of *Saccharomyces cerevisiae* with selected variants supporting up to 11-fold higher production of isoprenoids. Taken together, our method extends on existing CRISPR technologies by facilitating efficient mutagenesis of hundreds of nucleotides in cognate genomic contexts.

1. Introduction

Directed evolution is a powerful approach and has been extensively used for functional analysis and optimization of DNA sequences, gene functions and protein structures (Buller et al., 2015; Jeschek et al., 2016). In order to employ directed evolution, both the strategy for creating mutant libraries, as well as the assays for screening and selection of improved variants in high throughput, need to be considered (DeLoache et al., 2015; Körfer et al., 2016; Wong et al., 2004). Nowadays, methods for generation of sequence-variant libraries often include custom-designed DNA oligo library pools, multiplexed oligo assembly, or more or less randomized mutagenesis of template DNA fragments (Mutalik et al., 2013; Plesa et al., 2018; Yang et al., 2017). Screening and selection assays based on conditional growth, genetically-encoded biosensors, and reporter gene activities, has recently been applied to complement the efficiency of library generation (Lee et al., 2013; Mundhada et al., 2016; Raman et al., 2014; Skjoedt et al., 2016). Irrespective of the origin of input library and the screening assays applied, when aiming to infer functional consequences of regulatory or genic DNA variants, it is important to acknowledge the impact of gene dosage and the genomic context of the gene encoding the protein of interest, as both overexpression and epigenetic regulation can impact downstream regulation arising from genetic variants when expressed from non-native plasmid-based cassettes or when using heterologous promoters (Gibson et al., 2013). For this reason, targeted genome integration of diversified templates should be the preferred tactic for directed evolution. However, in order to harness the power of directed evolution in defined genomic contexts it is important to evaluate the capacity for living cells to take up and genomically integrate large, heterologous DNA fragments.

Pioneering studies on integration of linear fragments with homologous ends to chromosome integration sites was performed in yeast more than three decades ago (Orr-Weaver et al., 1981; Resnick and Martin, 1976; Szostak et al., 1983). Since then, the use of inducible endonucleases targeting defined genomic loci has further developed and improved both HR-based in vivo assembly of DNA fragments and methods for genome integration of heterologous DNA in both yeast and mammalian cells (Kuijpers et al., 2013; Rouet et al., 1994; Storici et al., 2001; Storici and Resnick, 2003). As such, the use of endonucleases has been reported to increase integration efficiencies by several orders of magnitude (Storici and Resnick, 2003). More recently, the use of the type II CRISPR/Cas9 system from *Streptococcus pyogenes* has gained broad adoption for RNA-guided gene editing and genome engineering in a wide range of organisms (Cong et al., 2013; DiCarlo et al., 2013; Jakočiūnas et al., 2017, 2016; Li et al., 2015). Cas9 is an RNA-guided endonuclease that introduces double-stranded breaks (DSBs) at almost any target DNA locus where a protospacer-adjacent motif (PAM; NGG) is present, yielding several hundred thousand genomic target sites in an

* Corresponding author.
E-mail address: mije@biosustain.dtu.dk (M.K. Jensen).

https://doi.org/10.1016/j.ymben.2018.07.001
Received 15 March 2018; Received in revised form 4 July 2018; Accepted 4 July 2018
Available online 05 July 2018

average eukaryotic genome (DiCarlo et al., 2013). With the ability to target and integrate heterologous DNA fragments at high efficiencies by way of homologous recombination, several studies have investigated directed evolution and sequence-function relationships of regulatory and genic regions in genomic contexts, albeit using short oligonucleotides as repair donors (Barbieri et al., 2017; Findlay et al., 2014; Garst et al., 2017; Storici et al., 2001; Wang et al., 2009). In addition, larger mutagenized DNA fragments were in vivo assembled and integrated into the *S. cerevisiae* genome by employing Cas9 and a selection marker (Ryan et al., 2014).

Here we employ CRISPR/Cas9 for introduction of DSB at genomic targets and native homology-directed repair to integrate mutagenised 300–600 bp linear DNA donors to demonstrate a simple, yet efficient, method for directed evolution of enzymes in native genomic contexts. We show that the method works for targeting single and double genomic loci with equal efficiency, and also show that even mutation distribution along the engineered fragment is maintained for the genomic integrants. Finally, we apply the method for directed evolution of two key metabolic enzymes in the mevalonate pathway of *S. cerevisiae*, thereby improving isoprenoid production by 11-fold. The method is called CasPER, short for Cas9-mediated protein evolution reaction.

2. Materials and methods

2.1. Strains, plasmids and media

The yeast strains used here were isogenic to CEN.PK2-1C. Strains and plasmids are listed in Supplementary Tables S1 and S2, respectively. Yeast cells were grown in complete medium (YPD) with 2% glucose and synthetic complete (SC) from Sigma, supplemented with 2% glucose. *E. coli* strains were propagated in LB medium supplemented with 200 mg of ampicillin.

All primer names and sequences are listed in Supplementary Table S3.

2.2. Selection of gRNAs and plasmid construction

To select for specific gRNAs targeting *BTS1*, *HMG2, ERG12* and *ERG20* we used the online CRISPy-web tool (Ronda et al., 2014). Selected gRNA sequences corresponded as follows: *BTS1* – TAGCTGCGA TACAAGTTGCA; *HMG2* – TGCTAGACATCTTCCCGGAT; *ERG12* – GTT AATAGGATCTAATGACT and *ERG20* – CTATGAAGGAGTGCTTCTTT. All gRNAs were expressed under SNR52 promoter and SUP4 terminator, and the expression cassettes were USER cloned into cloning vector pTAJAK-96 as previously described (Jakočiūnas et al., 2015a, 2015b). To construct the Cas9 expression plasmid, first centromeric pRS414 shuttle vector with the tryptophan selection marker (Sikorski and Hieter, 1989) was amplified with primers TJOS-167F and TJOS-167R to create USER cloning-compatible ends. Second, Cas9 expression cassette (TEF1 promoter, Cas9 and CYC1 terminator) was amplified from plasmid ID43802 (Addgene) using primers TJOS-168F and TJOS-168R containing compatible USER cloning ends. Third, TEF1-Cas9-CYC1 was USER cloned into previously prepared pRS414 shuttle vector labelled pTAJAK-162. To create yeast strain TC-146 expressing Cas9, pTAJAK-162 was transformed in to CEN.PK2-1C.

The carotenoid expressing strain was constructed as described previously (Ronda et al., 2015), herein labelled TC-100.

To reengineer the yeast strains with *ERG12* and *ERG20* mutant variants, first, *ERG12–71* and *ERG20–53* mutant variants were amplified with primers TJOS-127F and TJOS-127R; TJOS-126F and TJOS-126R respectively. Second, strain TC-146 was transformed either with a single gRNA expression plasmid (targeting *ERG12*: pTAJAK-118 or ERG20: pTAJAK-113) or with double gRNA expression plasmid (*ERG12* + *ERG20*: pTAJAK-187) and previously amplified *ERG12–71* - *ERG20–53* mutant variants, which served as donor templates for integration through homologous recombination and replacement of original *ERG12* or *ERG20* or both. Newly constructed strains (TC-150: ERG12::ERG12–71; TC-151: *ERG20::ERG20–53*; TC-152: ERG12::ERG12–71, *ERG20::ERG20–53*) were used further to integrate carotenoid genes and measure carotenoid production.

2.3. Mutagenesis and library preparation

First, templates for the error-prone PCR (epPCR) of the genes to be engineered were designed with the PAM site mutated to prevent cutting of the donor DNA, and subsequently ordered as biobricks (Integrated DNA Technologies). The PAM site was mutated in such a way that the obtained mutation would not change the native amino acid composition. Second, each biobrick served as a template for epPCR. The epPCR was performed with the GeneMorph II Random Mutagenesis kit (Agilent Technologies) using primers: *BTS1* – TJOS-325F and TJOS-325R; *HMG2* – TJOS-323F and TJOS-323R; *ERG12* – TJOS-124F and TJOS-124R; *ERG20* – TJOS-123F and TJOS-123R. Five consecutive rounds of epPCR were performed by transferring 50–100 ng of DNA to the next round of epPCR. The epPCR conditions were as follows:

95 °C–2 min .–30 × [95 °C–30 *s* – 60 °C–30s – 72 °C–1 min .]
–72 °C–10 min

After 5 rounds of epPCR, the resulting DNA fragments (*BTS1* or *HMG2*) were pooled and purified by ethanol precipitation adding 3 volumes of 99% ethanol and 1/10 vol of 3 M sodium acetate. As for *ERG12* and *ERG20*, after each round of epPCR, the resulting DNA was purified by Nucleospin PCR cleanup kit (Macherey-Nagel). Further, the DNA was treated with USER enzyme (New England Biolabs) according to the manufacturer's instructions. Three USER treatment reactions were performed with each purified epPCR reaction and transformed into *E. coli* competent cells together with pTAJAK-71 (Jessop-Fabre et al., 2016) previously digested with *Asi*SI (Thermo Fisher) and *Nb. Bsm*I (New England Biolabs) to create mutagenesis libraries. *E. coli* transformants were plated on LB supplemented with ampicillin and grown over night at 37 °C. Bacterial colonies were washed from the plates and grown for plasmid extraction. Plasmids were extracted by using Maxi prep kit (Qiagen).

2.4. Genome integration of mutagenized fragments

To integrate mutagenized fragments into the genome, yeast strain TC-100 or TC-3 harbouring Cas9 expression vector was used. Transformation with electroporation method (Wu and Letchworth, 2004) was performed with the following amounts of DNA:

5 picomoles of purified mutagenized fragments from each of the five epPCR's (25 picomoles in total), for multiplex targeting 25 picomoles of equally mixed *BTS1* mutagenized fragments and 25 picomoles of equally mixed *HMG2* mutagenized fragments (50 picomoles in total); for *ERG12* and *ERG20* engineering, each mutagenesis library was digested with *Pvu*II and fragments gel purified (25 picomoles in total for each *ERG12* and *ERG20* were used); 10 µg of single or double gRNA expression vector was used.

Mixed DNA needed for transformation was ethanol precipitated and resuspended in 5 µL of water. BioRad MicroPulser electroporator with settings for Fungi was used.

After transformation, cells were plated on SC (Sigma) plates or propagated in liquid media by selecting for Cas9 and gRNA expression vectors. Cells were propagated at 30 °C for 2–4 days.

2.5. Transformation efficiency quantification

Transformation efficiencies were quantified by plating dilution series on selective media and calculation the number of colonies (Supplementary Fig. S1).

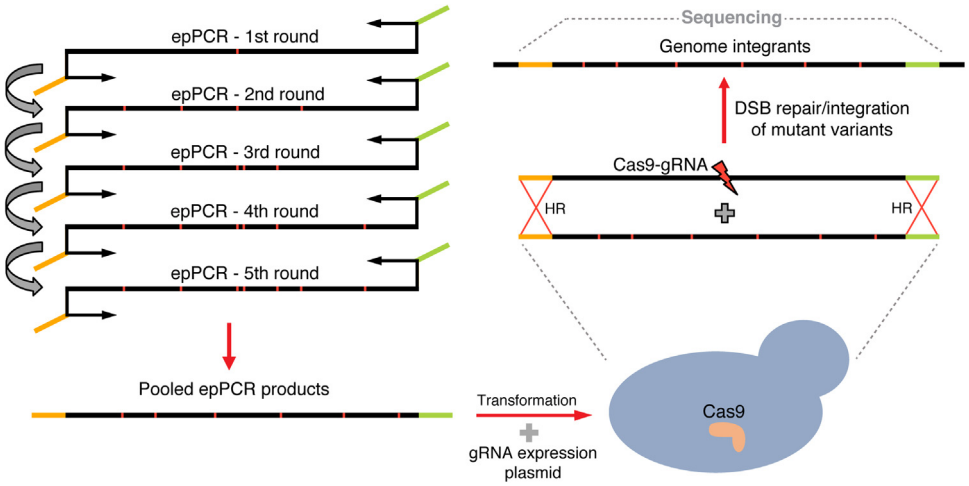

Fig. 1. Schematic overview of experimental setup of CasPER method. CasPER is based on generation of mutagenized linear DNA fragments through repetitive rounds of error-prone PCR (epPCR) and a standard yeast transformation together with gRNA expression plasmid for genome integration.

2.6. Sequencing and analysis

To characterize the method, both *BTS1* and *HMG2* mutagenized fragment libraries and genome integrants were sequenced. After propagation of yeast transformants in liquid media, the DNA from whole cultures was extracted using ZR Fungal/Bacterial DNA miniprep kit (Zymo Research). Both mutagenized fragments and genomic DNA were used to generate amplicons of *BTS1* and *HMG2* as described by Lee and colleagues (Lee et al., 2015) with the following primers: TJOS-339F and TJOS-339R; TJOS-340F and TJOS-340R respectively. Amplicons were sequenced by Illumina MiSeq using Nextera DNA library prep kit (Illumina).

Initially raw fastq files where processed using the r package DADA2 to infer exact amplicon variants and their abundance (Callahan et al., 2016). All additional analysis was performed in Jupyter-lab sing python 3.6.5 (Python Software Foundation, https://www.python.org/) and BioPython (Cock et al., 2009), Pandas (McKinney et al., 2010), Matplotlib (Hunter, 2007), numpy (Oliphant, 2015), mpmath (Johansson, 2013) and SciPy (Jones et al., 2001).

The analysis notebook and all raw data is available at https://data.mendeley.com/datasets/wkd5wp9zwk/1.

In brief the analysis consists of the following: identify type and number of mutants in detected variants compared to wild type and visualize data in a number of ways. Count mutants per position and visualize. Count specific mutation per position and visualize.

Count integration rate.

To evaluate if theres is a mutation frequency bias around the PAM site we did the following: Fit a negative binomial (nb) to a histogram representation of the mutation frequency per position along the amplicon. Then use the nb model to calculate how often you would see a mutation frequency equal to, or higher, than the frequency observed around the pam site (median of 5 bp window).

Furthermore, selected transformants from *BTS1, HMG2, ERG12* and *ERG20* targets were Sanger sequenced with the following primers: TJOS-15F and TJOS-15R (*BTS1*); TJOS-371F and TJOS-371R (*HMG2*); TJOS-127F and TJOS-127R (*ERG12*); TJOS-126F and TJOS-126R (*ERG20*).

2.7. Carotenoid extraction and quantification

Five mL cultures were inoculated and grown at 30 °C for 72 h and an equal amount of each culture (based on OD) was taken for further analysis. Cells were collected by centrifugation and cell walls broken using 0,5 mm glass beads (Sigma) and Precelllys 24 homogenizer (Bertin Instruments). Intracellular carotenoids were extracted with hexane.

In the pre-screen, carotenoid measurements were performed by measuring the absorption at 449 nm with Implen spectrophotometer (Implen). A standard curve was determined by measuring known concentrations of β-carotene and the following formula was used to quantify carotenoids in the extracts (Verwaal et al., 2007):

$$\text{Concentration of carotenoids} = 25{,}509 \times A_{449\,\text{nm}} - 0{,}5809.$$

For selected mutants, carotenoids (β-carotene and lycopene) were quantified by HPLC. Diluted culture supernatants were held in the autosampler at a temperature of 5 °C prior to analysis. The sample volume injected was 10 µL and the separation was performed using a Supelco Discovery HS F5-3 column (Sigma Aldrich), with a particle size of 3 µm, i.d. of 2.1 mm and a length of 15 cm. The column oven temperature was set to 30 °C. Gradient elution was performed at a constant flow rate of 0.7 mL/min using water with formic acid (1%) (eluent A) and acetonitrile (eluent B). The following gradient was applied: The initial isocratic step with 25% eluent B (2 min) was followed by an increase to 90% within 2 min and kept for another 6.5 min. The column was returned to initial conditions by decreasing eluent B for approx. 1 min to 25% and remained at this value for another 2 min. UV detection at 450 nm was applied for the quantification of β-carotene and lycopene.

2.8. The methodological workflow

CasPER can be divided into the following major steps (Fig. 1). First, a decision is made which DNA fragment is to be mutated. In our test case we initially mutagenized fragments of 300-bp and 600-bp. Second, a suitable gRNA is designed for the genomic target site(s) of interest by the use of CRISPy (Jakočiūnas et al., 2015a), or similar gRNA design tools, and a plasmid carrying gRNA expression cassette(s) for the chosen PAM(s) is constructed. Third, 300-bp and/or 600-bp PAM-deficient DNA templates for mutagenesis are synthesized or amplified from the host genome and processed through multiple rounds of error-prone PCR (epPCR). The number of PCR rounds applied is reflected in the number of mutations in the DNA fragment of interest (Cherry et al., 1999). The epPCR is performed with primers containing 52-bp overhangs homologous to the genomic loci upstream and downstream of the Cas9-mediated double-strand break (DSB). Overhangs may also include a restriction site or a motif for cloning and generation of plasmid-based mutagenesis library useful for propagation in *E. coli* and further use. Fourth, mutagenized linear fragments together with the plasmid carrying the gRNA expression cassette are co-transformed by electroporation into the organism of interest.

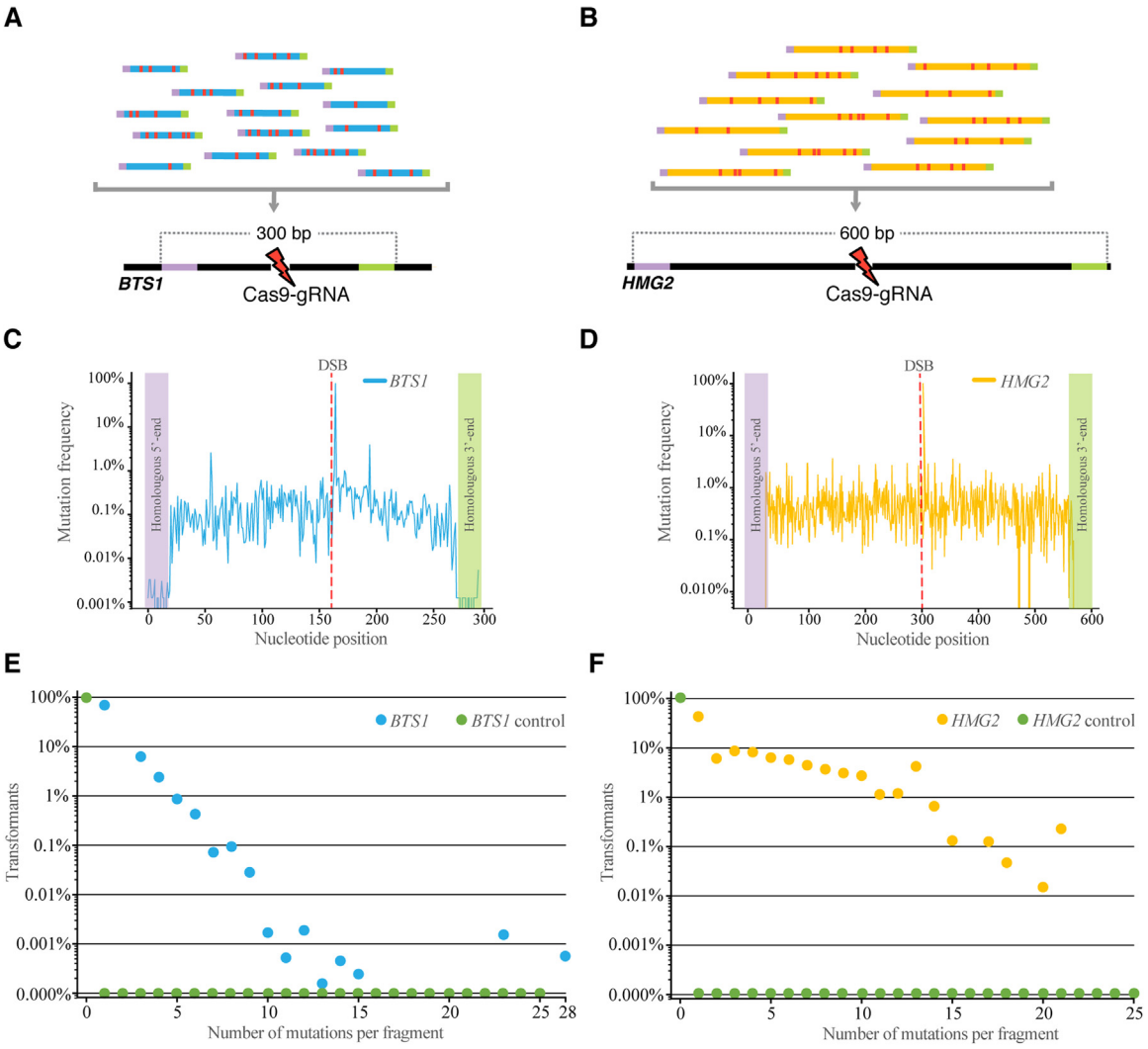

Fig. 2. Characterization of CasPER for directed evolution of genes. (**A**) A schematic overview of experimental setup for engineering of *BTS1,* 300 bp fragment. (**B**) A schematic overview of experimental setup for engineering of *HMG2,* 600 bp fragment. (**C**) Mutation landscape across 300-bp *BTS1* fragment. Frequencies of mutations for each nucleotide position are shown. Red line shows the position of double strand break (DSB) induced by Cas9. Purple and green rectangles indicate unmutated homologous ends that were introduced as primers overhangs during epPCR for the efficient integration to the homologous genomic site. (**D**) Mutation landscape across 600-bp *HMG2* fragment. Frequencies of mutations for each nucleotide position are shown. Red line shows the position of double strand break (DSB) induced by Cas9. Purple and green rectangles indicate unmutated homologous ends that were introduced as primers overhangs during epPCR for the efficient integration to the homologous genomic site. (**E**) Number of mutations per 300 bp *BTS1* fragment is shown in percentage of all sequenced transformants. (**F**) Number of mutations per 600-bp *HMG2* fragment is shown in percentage of all sequenced transformants.

3. Results

3.1. Characterization of CasPER for mutagenesis of genomic loci

As a proof-of-principle we applied the method in the yeast *S. cerevisiae* in order to engineer a 300-bp catalytic domain of the non-essential gene encoding geranylgeranyl diphosphate synthase (*BTS1;* Fig. 2A) and a 600-bp region of non-essential gene encoding HMG-CoA reductase (*HMG2;* Fig. 2B). To characterize the size and sequence composition of integrated epPCR-derived donor fragments, we sequenced genomic amplicons from approximately 2000 colonies by next-generation sequencing (NGS). This number of colonies was targeted to get enough reads per colony-amplicon to overcome well-known NGS error rates (Huang et al., 2012; Manley et al., 2016; Raymond et al., 2017; Schirmer et al., 2015). The generated NGS data was analysed to characterize the mutation landscapes of the genomic target loci and to estimate the number of genome integrants at BTS1 and HMG2 genomic loci compared to the mutation landscapes arising without the use of Cas9-mediated DNA double-strand break (i.e. no gRNAs introduced).

First, since DNA double-strand breaks in yeast are primarily repaired by homologous recombination, the degree of sequence homology at the termini of linear donor fragments is known to affect the frequency of insertional repair (Haviv-Chesner et al., 2007). To evaluate if the method would support integration of linear fragments mutagenized along the full length of the donor fragments, we mapped all sequenced mutations across the 300-bp and 600-bp region and calculated the frequency at which mutations occurred in a given position (Fig. 2C-D; Supplementary Fig. S2). From this analysis, we observed that mutations distribute evenly across both the 300-bp and 600-bp regions with no overall bias towards higher mutation frequencies in proximity to the DSB site (Fig. 2D-C; Supplementary Fig. S3). This is an important observation when designing the genomic region to be targeted (e.g. evolution of larger catalytic domains of enzymes).

Furthermore, sequencing revealed that, following standard transformation procedure of *S. cerevisiae,* 98–99% Cas9-mediated integration efficiency was observed for both *BTS1* and *HMG2* (i.e. all sequenced colonies had at least one mutation integrated into the genomic site) (Supplementary Table S4). Sequenced variants had up to twenty eight

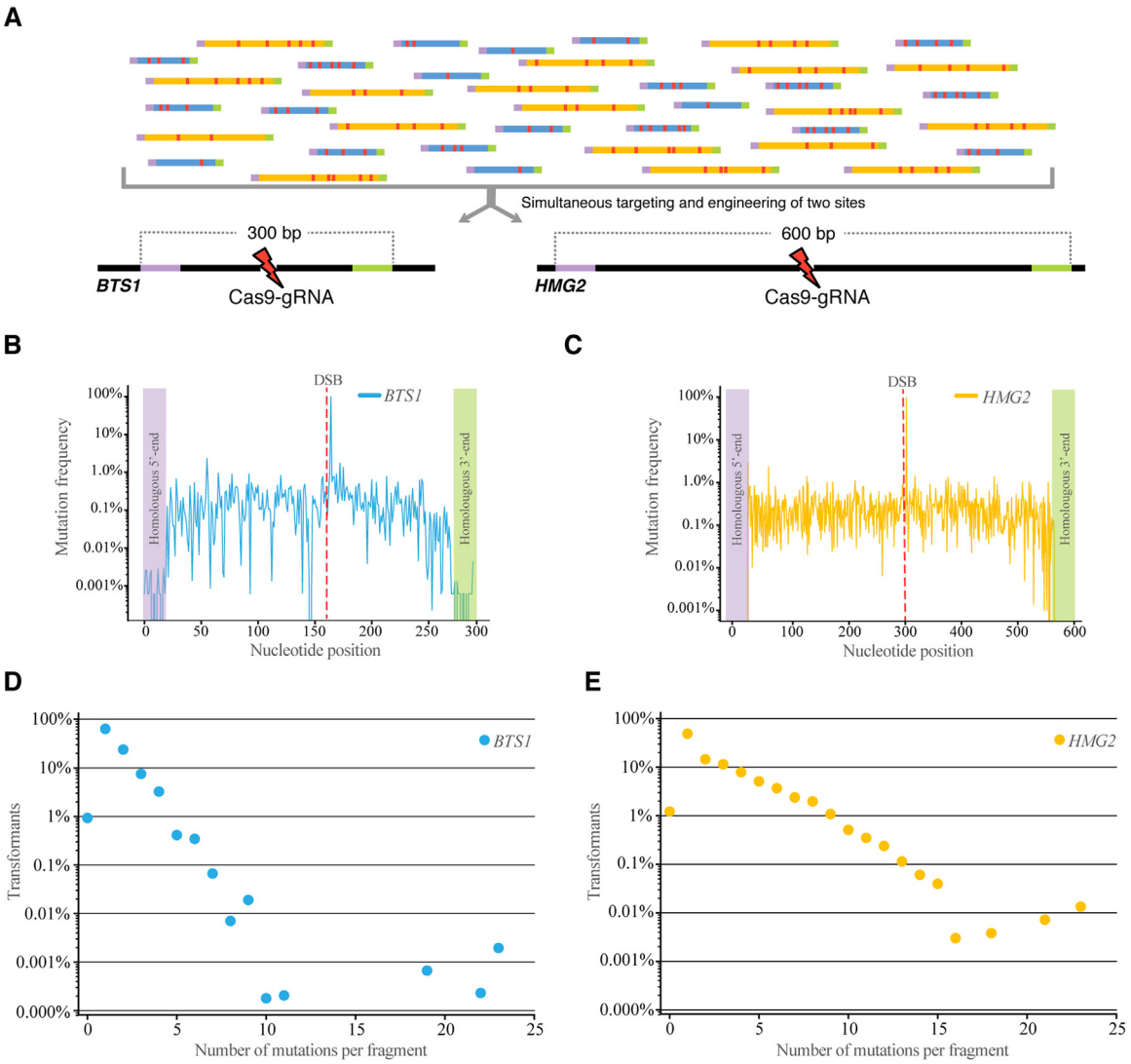

Fig. 3. Characterization of CasPER for directed evolution of genes in multiplex manner. (**A**) A schematic overview of experimental setup for engineering of *BTS1* (300 bp) and *HMG2* (600 bp) simultaneously. (**B**) Mutation landscape across 300-bp *BTS1* fragment when targeted simultaneously with *HMG2*. Frequencies of mutations for each nucleotide position are shown. Red line shows the position of double strand break (DSB) induced by Cas9. Purple and green rectangles indicate unmutated homologous ends that were introduced as primers overhangs during epPCR for the efficient integration to the homologous genomic site. (**C**) Mutation landscape across 600-bp *HMG2* fragment when targeted simultaneously with *BTS1*. Frequencies of mutations for each nucleotide position are shown. Red line shows the position of double strand break (DSB) induced by Cas9. Purple and green rectangles indicate unmutated homologous ends that were introduced as primers overhangs during epPCR for the efficient integration to the homologous genomic site. (**D**) Number of mutations per 300 bp *BTS1* fragment is shown in percentage of all sequenced transformants. (**E**) Number of mutations per 600-bp *HMG2* fragment is shown in percentage of all sequenced transformants.

mutations when targeting *BTS1* and up to twenty one mutations when targeting *HMG2*, with a peak of one mutation per both 300-bp long fragment and 600-bp long fragment (Fig. 2E-F). Further, as a control experiment for the DSB-induced genome engineering efficiency, we transformed the same amount of mutagenized fragments into yeast expressing Cas9, but without co-transforming the gRNA expression plasmid. Here, no mutagenized fragments were observed for *BTS1* and *HMG2*, confirming that DSB caused by Cas9 is an essential feature of the method (Fig. 2E-F).

Next, to characterize the efficiency of CasPER, we estimated the total number of possible variants by calculating transformation efficiency when *BTS1* and *HMG2* are targeted. From this estimate we obtained almost 2,4 million transformants when targeting *BTS1* and almost 1,6 million transformants when targeting *HMG2* (Supplementary Fig. S1). Hence, since integration efficiency at both tested loci is 98–99% (Supplementary Table S4), the number of transformants obtained is believed to be a reasonable proxy for estimating the number of genomically integrated mutant fragments.

Biological systems are prone to genetic epistasis and synthetic lethality (Mackay, 2014; Phillips, 2008). In order to uncover such relationships and engineer genomes in higher throughput, genome engineering methods should preferably support directed evolution of more than one genomic locus per transformation. For this purpose, we tested whether CasPER can enable the integration of mutagenized linear fragments into two genomic loci in a single transformation. To test this, we designed an experiment in which *BTS1* and *HMG2* were targeted simultaneously using the same mutagenized 300-bp and 600-bp DNA fragments as in single targeting experiment described previously (Fig. 3A). As for the characterization of the single locus edits (Fig. 2C-F), we performed sequencing analysis to characterize the mutation frequency and the mutation landscapes of the two genomic target loci, when targeted in a single transformation. From this we observed that the integration efficiencies of mutagenized fragments were were identical to the efficiencies observed for singleplex targeting, namely

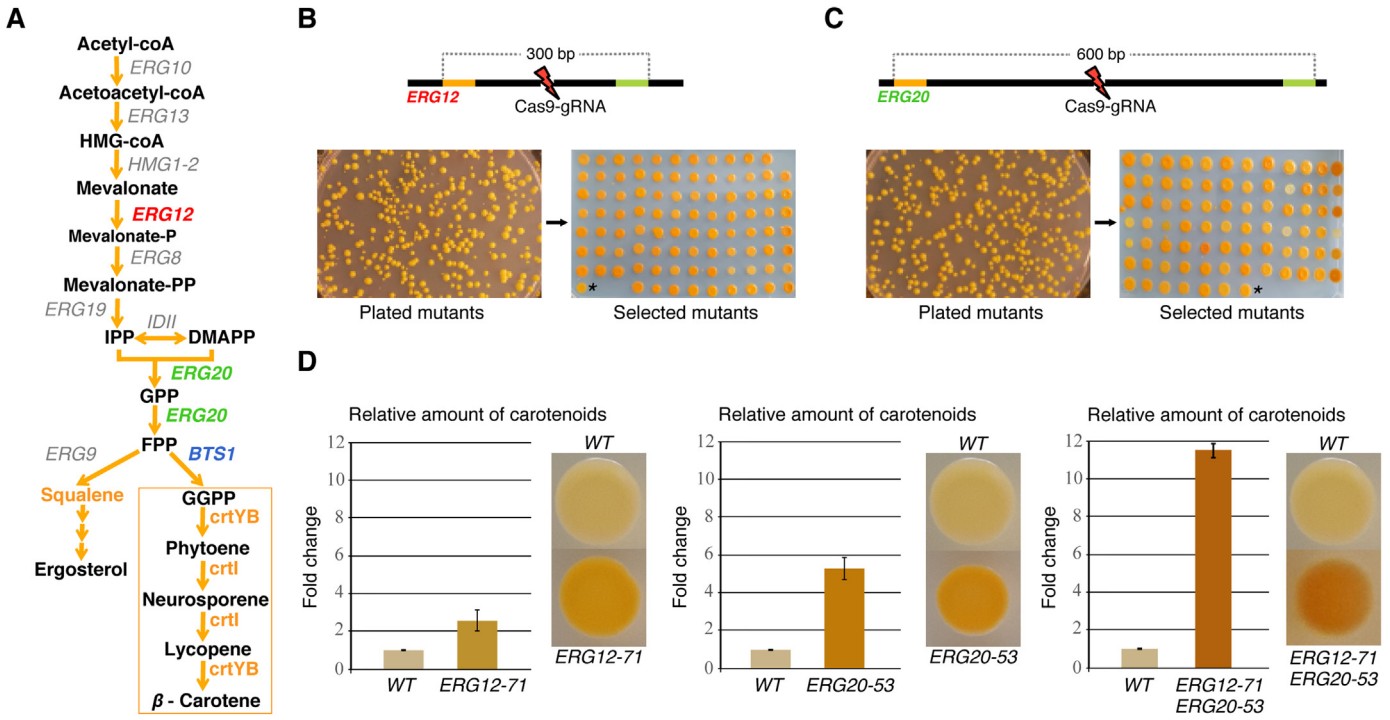

Fig. 4. Application of CasPER for engineering key enzymes in ergosterol biosynthesis pathway. (**A**) Schematic outline of the yeast mevalonate pathway: engineered enzymes are depicted in red (*ERG12*), green (*ERG20*) and blue (*BTS1*). Carotenoid biosynthetic pathway is shown in orange rectangle and pathway genes written in orange. (**B**) Overview of engineering and screening process for *ERG12*. (*) – marks WT colony. (**C**) Overview of engineering and screening process of *ERG20*. (*) – marks WT colony. (**D**) Relative carotenoid measurements for wild-type and selected single and double mutants are presented together with visual colour phenotype of the cells.

98% for *BTS1* and 99% for *HMG2* (Supplementary Table S4). As for the mutation landscape along the 5' − 3' end of single donor fragments, we again mapped all sequenced mutations across the 300-bp and 600-bp region when both loci were targeted simultaneously (Fig. 3B–C; Supplementary Fig. S4). Similar to our analysis of single genomic loci, the mutation landscapes across the 300-bp and 600-bp regions showed that mutations are evenly distributed without bias towards the DSB (Fig. 3B–C; Supplementary Fig. S5). Additionally, sequenced variants had up to twenty three mutations each when targeting both *BTS1* and *HMG2*, with a peak of one mutation for both 300-bp long fragment and 600-bp long fragment (Fig. 3D–E).

To characterise the efficiency of CasPER in multiplex manner, we estimated the total number of possible variants by calculating transformation efficiency when *BTS1* and *HMG2* is targeted simultaneously. We obtained over 1 million transformats (Supplementary Fig. S1) with integration efficiency of 98–99% (Supplementary Table S4), similarly to what was observed for single site engineering.

To summarize, these results show that CasPER allows for efficient integration of large donor fragments into both single and multiple genomic targets. Furthermore, the method is tolerant to genomic integration of donor fragments with multiple mutations without observed biases along the donor regions. In comparison to other existing genome engineering methods, CasPER is specifically useful when targeting larger genomic regions, without sacrificing engineering efficiency and cost.

3.2. Application of CasPER for engineering key enzymes in yeast mevalonate biosynthesis pathway

In order to investigate the potential of CasPER for directed evolution of essential genes in their genomic contexts, we targeted two enzymes in the mevalonate pathway, namely *ERG12* and *ERG20* (Fig. 4A). Control of metabolic flux through the mevalonate pathway is of paramount importance for prevention of various diseases (Akula et al., 2016; Frenkel et al., 2000), and for the development of biobased production of pharmaceuticals, food additives, fuels, cosmetics and others (Ye et al., 2016). In yeast, *ERG12* encodes the mevalonate kinase, responsible for phosphorylation of mevalonate, whereas *ERG20* encodes farnesyl pyrophosphate synthase, which catalyzes the formation of both farnesyl pyrophosphate (FPP) and geranyl pyrophosphate (GPP), units for sterol and isoprenoid biosynthesis (Chambon et al., 1991; Oulmouden and Karst, 1990). It has been shown that FPP and GPP can feedback inhibit binding of ATP to human mevalonate kinase, and hence limit the production through the pathway further (Ruff et al., 2014).

We applied CasPER to mutate the 300-bp kinase catalytic domain of *ERG12*, and the 600-bp catalytic domain of *ERG20* (Fig. 4B, C). In order to efficiently screen and select enzyme variants supporting improved pathway flux, we introduced a metabolic sink downstream from the mevalonate and FPP pools, comprised of two heterologous carotenoid biosynthetic genes *crtYB* and *crtI* from yeast *Xanthophyllomyces dendrorhous* (Fig. 4A). When carotenoids are produced through this heterologous pathway, yeast transformants form yellow to orange colonies, depending on the amount of carotenoids produced (Özaydın et al., 2013; Verwaal et al., 2007). Hence, this pathway provides a facile visual screening platform to complement the high efficiency of CasPER. Following transformation, approximately 60,000 colonies from each of the two library transformations were plated. Next, approximately 200 colonies displaying a darker orange phenotype than the wild-type reference strain were selected for further characterization (Fig. 4B, C). Following colony replication and repeated scoring of the phenotype, a pre-screen using spectrophotometric measurements of carotenoid extracts showed 1.5–3-fold higher carotenoid levels in 24 *ERG12* and 11 *ERG20* mutants (Supplementary Fig. S6). One of each of the *ERG12* and *ERG20* mutants displaying darker orange phenotypes were selected for further analyses. First, the selected mutants, *ERG12–71* and *ERG20–53*, were re-introduced into a wild-type strain replacing the native *ERG12* and *ERG20* genes. Both single and double mutant strains were

A

Alignment of mevalonate kinase (*ERG12*)

Saccharomyces cerevisiae	178	NDKHIVNQ	W	AFIGEKCIHGTPSGIDNAVATYGNALLFEKDSHNGTINTNNFKFLD	232
Schizosaccharomyces pombe	171	KALALIEA	W	SFLGECCIHGTPSGIDNAVATNGGLIAFRKATAHQSAMK----EFLKPKDT	226
Arabidopsis thaliana	177	ETNLELLNK	W	AFEGEKIIHGKPSGIDNTVSAYGNMIKF----CSGEIT----RLQSNM	226
Vitis vinifera	179	ENELELVNK	W	AFEGEKIIHGKPSGIDNSVSTFGNMIKFKSGSLT--------RLSSNM	228
Mus musculus	180	EEDLKSINK	W	AFEGERVIHGNPSGVDNAVSTWGGALRFQQ----GTMSS-----LKSLPSL	238
Homo sapiens	179	KEDLELINK	W	AFQGERMIHGNPSGVDNAVSTWGGALRY----HQGKISS-----LKRSPA	229
S. cerevisiae (mutant #71)	178	NDKHIVNQ	R	AFIGEKCIHGTPSGIDNAVATYGNALLFEKDSHNGTINTNNFKFLD	232

B

Alignment of farnesyl pyrophosphate synthase (*ERG20*)

Saccharomyces cerevisiae	68	NKTVEQLGQEEYEKVAILGWCIELLQA	Y	FLVADDMMDKSITRRGQPCWYKVPEVGEIAIN	127
Schizosaccharomyces pombe	66	G---HPLDEAAYMKAAVLGWMVELLQS	F	FLIADDIMDASKTRRGQPCWYLMPGVGNIAIN	122
Arabidopsis thaliana	104	KQGNDLTEQEVFLSCALGWCIEWLQA	Y	FLVLDDIMDNSVTRRGQPCWFRVPQVGMVAIND	163
Vitis vinifera	64	R---QLTDDEVFLACVLGLCIEWLQA	Y	FLVLDDIMDNSHTRRGQPCWFRVPKVGMIAAND	120
Mus musculus	74	PKKQD--AESLQRALTVGWCVELLQA	F	FLVSDDIMDSSLTRRGQICWYQKPGIGLDAIND	131
Homo sapiens	72	VEPRKQDADSLQRAWTVGWCVELLQA	F	FLVADDIMDSSLTRRGQTCWYQKPGVGLDAI	129
S. cerevisiae (mutant #53)	68	NKTVEQLGQEEYEKVAILGWCIELLQA	S	FLVADDMMDKSITRRGQPCWYKVPEVGEIAIN	127

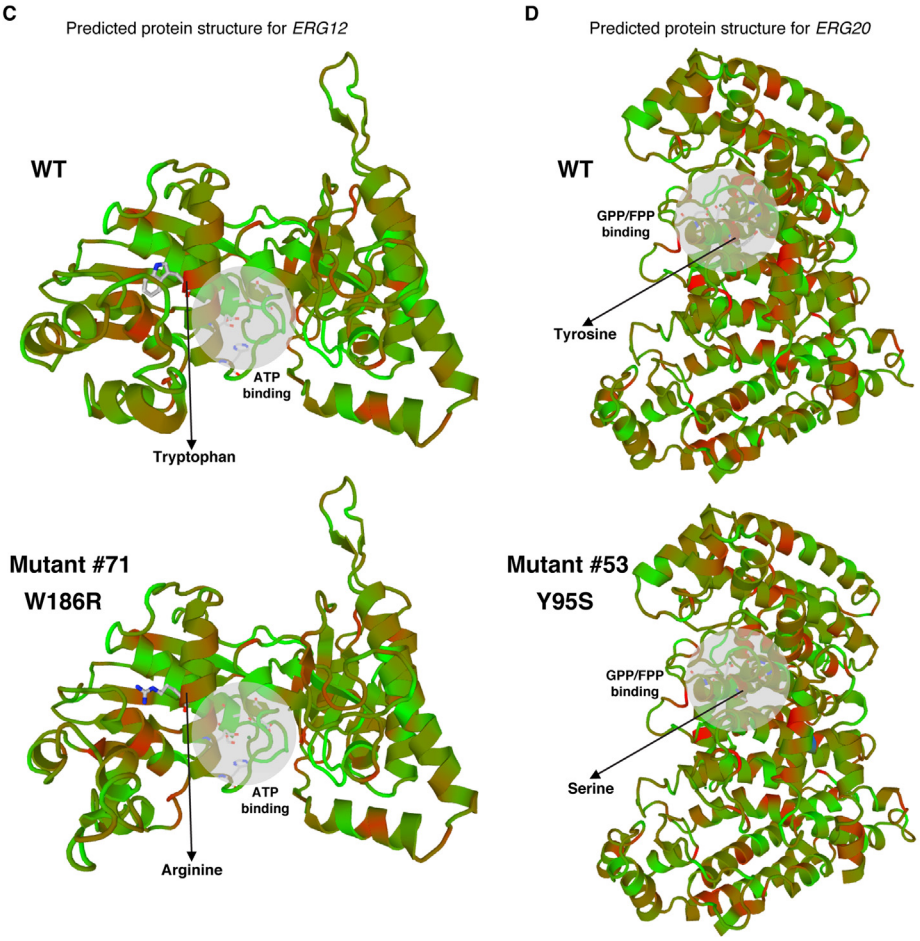

Fig. 5. Mutant analysis of yeast *ERG12* and *ERG20*. (**A**) Amino acid sequence alignment of *S. cerevisiae WT ERG12, ERG12–71* and mevalonate kinases in species form plant, fungi and mammalian kingdoms. Red box shows conserved tryptophan at position 186 through different species and mutant 71 with tryptophan mutated to arginine (**B**) Amino acid sequence alignment of *S. cerevisiae WT ERG20, ERG20–53* and farnesyl pyrophosphate synthases in species form plant, fungi and mammalian kingdoms. Red box shows conserved tyrosine at position 95 through different species and mutant 53 with tyrosine mutated to serine (**C**) Predicted 3D protein structure of *WT ERG12* and *ERG12–71* is shown. Tryptophan position is depicted as well as ATP bound to WT *ERG12* is shown in upper panel. Arginine position is depicted as well as ATP bound to *ERG12–71* is shown in lower panel. (**D**) Predicted 3D protein structure of *WT ERG20* and *ERG20–53* is shown. Tyrosine position is depicted as well as GPP/FPP bound to *WT ERG20* is shown in upper panel. Serine position is depicted as well as GPP/FPP bound to *ERG20–53* is shown in lower panel.

constructed. Second, the carotenoid biosynthetic genes were introduced to the single and double mutant strains, and carotenoid production was quantified using HPLC. Carotenoid measurements showed an approx. 3-fold higher production with the *ERG12–71* mutant and a 5-fold higher production with the *ERG20–53* mutant, compared to the wild type reference strain expressing the carotenoid biosynthetic genes (Fig. 4D). In addition, the double *ERG12–71/ERG20–53* mutant strain showed an 11-fold higher production of carotenoids when compared to a wild-type reference strain (Fig. 4D).

Taken together, the application of CasPER for directed evolution of *ERG12* and *ERG20* allows for simple and robust uncovering of improved enzyme variants supporting increased flux through the feedback regulated isoprenoid pathway.

3.3. Mutant analysis of yeast mevalonate kinase and farnesyl pyrophosphate synthase

To investigate the structural basis for increased carotenoid levels supported by the *ERG12* and *ERG20* mutated enzymes we compared them to homologs from other organisms and positioned the mutants onto structural models. First, the sequence analysis of *ERG12–71* and *ERG20–53* showed a single amino acid change at position 186 from tryptophan to arginine for *ERG12*, and at position 95 from tyrosine to serine for *ERG20* (Supplementary Fig. S7). Second, we aligned the primary structure of the *ERG12* and *ERG20* mutants with homologs from mammals, plants, and other yeasts (Fig. 5A, B). Alignments showed that tryptophan in position 186 of *ERG12* is highly conserved throughout the kingdoms and can play a role in enzyme structure (Fig. 5A). On the other hand, tyrosine in position 95 of *ERG20* is

conserved in several fungi and plant species, yet different from the mammalian homologs, although, in mammals the tyrosine is replaced with phenylalanine (Fig. 5B). Next, we analysed if the mutations in *ERG12* and *ERG20* are causing structural changes to the enzymes. Since the crystal structures for yeast *ERG12* and *ERG20* are not available, we used SWISS-MODEL (Arnold et al., 2006; Biasini et al., 2014; Guex et al., 2009; Kiefer et al., 2009) for the alignment of protein structures based on known crystal structures from other organisms (Fig. 5C, D). For the *ERG12* structure alignment, the template 1kvk.1. A (Fu et al., 2002) was used, which resulted in 32.71% sequence identity and produced global model quality estimation (GMQE) of 0.6. For the *ERG20* structure alignment, the crystal structure of *Gallus gallus* FPP synthase was used as a template (1ubw.1. A)(Tarshis et al., 1996). This resulted in 47.65% sequence identity (GMQE = 0,77). We modelled 3D structures, both for wild-type *ERG12* and mutant *ERG12–71* (Fig. 5C) as well as for wild-type *ERG20* and *ERG20–53* mutant (Fig. 5D). The mutations in *ERG12* and *ERG20* do not visually influence the aligned 3D structures of both enzymes. Further, it is worth mentioning that both mutations in *ERG12* and *ERG20* are in close proximity to the ATP binding pocket of the mevalonate kinase (Fig. 5C) and the GPP/FPP binding pocket of farnesyl pyrophosphate synthase (Fig. 5D).

Finally, although our analysis does not provide a clear evidence to how the mutations in *ERG12* and *ERG20* influence activity or structure of the enzymes, it shows that the amino acids which were mutated are conserved and are situated in close proximity to the ligand binding pocket.

4. Discussion

In the last five years CRISPR/Cas9 has been used for genome engineering and re-programing in yeast with efficiencies reaching 100% (DiCarlo et al., 2013; Jakočiūnas et al., 2015a, b; Jensen et al., 2017; Mans et al., 2015; Ryan et al., 2014) in both single and multiplex manner, hence these advancements have revolutionised the way yeast genomes are engineered today. Beyond the existing technologies, CasPER is an efficient method, which combines the generation of large mutated DNA fragments by epPCR (or any other method for generating sequence diversity) and multiplex genome integration of these fragments by native HR machinery following targeted CRISPR/Cas9-mediated DNA DSBs.

Indeed, CasPER has proven to be a very efficient method, reaching nearly 100% efficiency of integration of large mutagenised DNA fragments in multiplex manner, and thereby allowing the generation of millions of mutants from a single standard transformation. Importantly, no significant bias in mutation frequency was observed along integrated donor fragments in relation to DSB proximity, allowing efficient and non-selective mutagenesis of large DNA fragments. As such, beyond the proven efficiency of introducing donor variant libraries up to 600 bp without the use of selection markers, this study also provides an in-depth demonstration of the mutational landscapes and frequencies which can be obtained surrounding CRISPR/Cas9-mediated DSB target sites in HR-proficient organisms like *S. cerevisiae*.

Further, in our study we hypothesized that CasPER can support the integration of large donor variants encoding enzyme catalytic domains or even whole coding regions of enzymes and other DNA regulatory elements. To test this, we applied CasPER for directed evolution of catalytic domains of two essential enzymes in the mevalonate pathway. The experiment uncovered mutants that were able to improve carotenoid production by 11-fold, as a proxy for changed metabolic flux through yeast's native mevalonate pathway. Though this study did not further characterize the biochemical activity of the evolved *ERG12* and *ERG20* mutants, it is anticipated that these variants will be subject to numerous follow-up studies and applications related to microbial production of fuels and chemicals derived from the mevalonate pathway.

Finally, at present, most of the CRISPR/Cas9-based directed evolution methods rely on integration of relative small (80–120 bp) DNA fragments (Barbieri et al., 2017; Garst et al., 2017; Guo et al., 2018) and may require time-consuming construction of donor variant libraries and/or relative costly DNA synthesis of diversified array-based oligos (Nyerges et al., 2018; Roy et al., 2018). With CasPER, multiplex genome engineering of larger genomic loci is now demonstrated at very high efficiencies along the full size of donor fragment lengths, and should thereby enable cost-effective, high-throughput and robust evolution studies of even complex multi-genic traits in any organism supporting genome integration of heterologous DNA by homologous recombination.

Acknowledgements

The authors would like to acknowledge Dushica Arsovska, Anna Koza, Alexandra Hoffmeyer and Pannipa Pornpitakpong for technical assistance in relation to NGS.

Funding

This work was funded by the Novo Nordisk Foundation.

Conflicts of interest

JDK has a financial interest in Amyris, Lygos, Demetrix, Constructive Biology, Maple Bio, and Napigen.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ymben.2018.07.001.

References

Akula, M.K., Shi, M., Jiang, Z., Foster, C.E., Miao, D., Li, A.S., Zhang, X., Gavin, R.M., Forde, S.D., Germain, G., Carpenter, S., Rosadini, C.V., Gritsman, K., Chae, J.J., Hampton, R., Silverman, N., Gravallese, E.M., Kagan, J.C., Fitzgerald, K.A., Kastner, D.L., Golenbock, D.T., Bergo, M.O., Wang, D., 2016. Control of the innate immune response by the mevalonate pathway. Nat. Immunol. 17, 922–929.

Arnold, K., Bordoli, L., Kopp, J., Schwede, T., 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22, 195–201.

Barbieri, E.M., Muir, P., Akhuetie-Oni, B.O., Yellman, C.M., Isaacs, F.J., 2017. Precise editing at DNA replication forks enables multiplex genome engineering in eukaryotes. Cell. https://doi.org/10.1016/j.cell.2017.10.034.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T.G., Bertoni, M., Bordoli, L., Schwede, T., 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res. 42, W252–W258.

Buller, A.R., Brinkmann-Chen, S., Romney, D.K., Herger, M., Murciano-Calles, J., Arnold, F.H., 2015. Directed evolution of the tryptophan synthase β-subunit for stand-alone function recapitulates allosteric activation. Proc. Natl. Acad. Sci. USA 112, 14599–14604.

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P., 2016. DADA2: high-resolution sample inference from Illumina amplicon data. Nat. Methods 13, 581–583.

Chambon, C., Ladeveze, V., Servouse, M., Blanchard, L., Javelot, C., Vladescu, B., Karst, F., 1991. Sterol pathway in yeast. Identification and properties of mutant strains defective in mevalonate diphosphate decarboxylase and farnesyl diphosphate synthetase. Lipids 26, 633–636.

Cherry, J.R., Lamsa, M.H., Schneider, P., Vind, J., Svendsen, A., Jones, A., Pedersen, A.H., 1999. Directed evolution of a fungal peroxidase. Nat. Biotechnol. 17, 379–384.

Cock, P.J.A., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., de Hoon, M.J.L., 2009. Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics 25, 1422–1423.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., 2013. Multiplex genome engineering using CRISPR/Cas systems. Science 339. https://doi.org/10.1126/science.1231143.

DeLoache, W.C., Russ, Z.N., Narcross, L., Gonzales, A.M., Martin, V.J.J., Dueber, J.E., 2015. An enzyme-coupled biosensor enables (S)-reticuline production in yeast from glucose. Nat. Chem. Biol. 11, 465–471.

DiCarlo, J.E., Norville, J.E., Mali, P., Rios, X., Aach, J., Church, G.M., 2013a. Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems. Nucleic Acids Res. 41, 4336–4343.

Findlay, G.M., Boyle, E.A., Hause, R.J., Klein, J.C., Shendure, J., 2014. Saturation editing

of genomic regions by multiplex homology-directed repair. Nature 513, 120–123.
Frenkel, J., Houten, S.M., Waterham, H.R., Wanders, R.J., Rijkers, G.T., Kimpen, J.L., Duran, R., Poll-The, B.T., Kuis, W., 2000. Mevalonate kinase deficiency and Dutch type periodic fever. Clin. Exp. Rheumatol. 18, 525–532.
Fu, Z., Wang, M., Potter, D., Miziorko, H.M., Kim, J.-J.P., 2002. The structure of a binary complex between a mammalian mevalonate kinase and ATP: insights into the reaction mechanism and human inherited disease. J. Biol. Chem. 277, 18134–18142.
Garst, A.D., Bassalo, M.C., Pines, G., Lynch, S.A., Halweg-Edwards, A.L., Liu, R., Liang, L., Wang, Z., Zeitoun, R., Alexander, W.G., Gill, R.T., 2017. Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. Nat. Biotechnol. 35, 48–55.
Gibson, T.J., Seiler, M., Veitia, R.A., 2013. The transience of transient overexpression. Nat. Methods 10, 715–721.
Guex, N., Peitsch, M.C., Schwede, T., 2009. Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. Electrophoresis 30, S162–S173.
Guo, X., Chavez, A., Tung, A., Chan, Y., Kaas, C., Yin, Y., Cecchi, R., Garnier, S.L., Kelsic, E.D., Schubert, M., DiCarlo, J.E., Collins, J.J., Church, G.M., 2018. High-throughput creation and functional profiling of DNA sequence variant libraries using CRISPR-Cas9 in yeast. Nat. Biotechnol. https://doi.org/10.1038/nbt.4147.
Haviv-Chesner, A., Kobayashi, Y., Gabriel, A., Kupiec, M., 2007. Capture of linear fragments at a double-strand break in yeast. Nucleic Acids Res. 35, 5192–5202.
Huang, W., Li, L., Myers, J.R., Marth, G.T., 2012. ART: a next-generation sequencing read simulator. Bioinformatics 28, 593–594.
Hunter, J.D., 2007. Matplotlib: a 2D Graphics Environment. Comput. Sci. Eng. 9, 90–95.
Jakočiūnas, T., Bonde, I., Herrgård, M., Harrison, S.J., Kristensen, M., Pedersen, L.E., Jensen, M.K., Keasling, J.D., 2015a. Multiplex metabolic pathway engineering using CRISPR/Cas9 in Saccharomyces cerevisiae. Metab. Eng. 28, 213–222.
Jakočiūnas, T., Rajkumar, A.S., Zhang, J., Arsovska, D., Rodriguez, A., Jendresen, C.B., Skjødt, M.L., Nielsen, A.T., Borodina, I., Jensen, M.K., Keasling, J.D., 2015b. CasEMBLR: Cas9-Facilitated Multiloci Genomic Integration of in Vivo Assembled DNA Parts in Saccharomyces cerevisiae. ACS Synth. Biol. 4, 1226–1234.
Jakočiūnas, T., Jensen, M.K., Keasling, J.D., 2017. System-level perturbations of cell metabolism using CRISPR/Cas9. Curr. Opin. Biotechnol. 46, 134–140.
Jakočiūnas, T., Jensen, M.K., Keasling, J.D., 2016. CRISPR/Cas9 advances engineering of microbial cell factories. Metab. Eng. 34, 44–59.
Jensen, E.D., Ferreira, R., Jakočiūnas, T., Arsovska, D., Zhang, J., Ding, L., Smith, J.D., David, F., Nielsen, J., Jensen, M.K., Keasling, J.D., 2017. Transcriptional reprogramming in yeast using dCas9 and combinatorial gRNA strategies. Microb. Cell Fact. 16, 46.
Jeschek, M., Reuter, R., Heinisch, T., Trindler, C., Klehr, J., Panke, S., Ward, T.R., 2016. Directed evolution of artificial metalloenzymes for in vivo metathesis. Nature 537, 661–665.
Jessop-Fabre, M.M., Jakočiūnas, T., Stovicek, V., Dai, Z., Jensen, M.K., Keasling, J.D., Borodina, I., 2016. EasyClone-MarkerFree: a vector toolkit for marker-less integration of genes into *Saccharomyces cerevisiae* via CRISPR-Cas9. Biotechnol. J. 11, 1110–1117.
Johansson, F., 2013. mpmath: a Python library for arbitrary-precision floating-point arithmetic (version 0.18) [WWW Document]. URL ⟨http://mpmath.org/⟩.
Jones E., Oliphant T., Peterson P., et al, 2001. SciPy: Open Source Scientific Tools for Python [WWW Document]. URL ⟨http://www.scipy.org/⟩.
Kiefer, F., Arnold, K., Kunzli, M., Bordoli, L., Schwede, T., 2009. The SWISS-model repository and associated resources. Nucleic Acids Res. 37, D387–D392.
Körfer, G., Pitzler, C., Vojcic, L., Martinez, R., Schwaneberg, U., 2016. In vitro flow cytometry-based screening platform for cellulase engineering. Sci. Rep. 6, 26128.
Kuijpers, N.G.A., Chroumpi, S., Vos, T., Solis-Escalante, D., Bosman, L., Pronk, J.T., Daran, J.-M., Daran-Lapujade, P., 2013. One-step assembly and targeted integration of multigene constructs assisted by the I-SceI meganuclease in Saccharomyces cerevisiae. FEMS Yeast Res. 13, 769–781.
Lee, J.S., Kallehauge, T.B., Pedersen, L.E., Kildegaard, H.F., 2015. Site-specific integration in CHO cells mediated by CRISPR/Cas9 and homology-directed DNA repair pathway. Sci. Rep. 5, 8572.
Lee, M.E., Aswani, A., Han, A.S., Tomlin, C.J., Dueber, J.E., 2013. Expression-level optimization of a multi-enzyme pathway in the absence of a high-throughput assay. Nucleic Acids Res. 41, 10668–10678.
Li, Y., Lin, Z., Huang, C., Zhang, Y., Wang, Z., Tang, Y.-J., Chen, T., Zhao, X., 2015. Metabolic engineering of Escherichia coli using CRISPR–Cas9 meditated genome editing. Metab. Eng. 31, 13–21.
Mackay, T.F.C., 2014. Epistasis and quantitative traits: using model organisms to study gene-gene interactions. Nat. Rev. Genet. 15, 22–33.
Manley, L.J., Ma, D., Levine, S.S., 2016. Monitoring Error Rates In Illumina Sequencing. J. Biomol. Tech. 27, 125–128.
Mans, R., van Rossum, H.M., Wijsman, M., Backx, A., Kuijpers, N.G.A., van den Broek, M., Daran-Lapujade, P., Pronk, J.T., van Maris, A.J.A., Daran, J.-M.G., 2015. CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in Saccharomyces cerevisiae. FEMS Yeast Res. 15. https://doi.org/10.1093/femsyr/fov004.
McKinney, W., 2010. Others. Data structures for statistical computing in python, In: Proceedings of the 9th Python in Science Conference. Austin, TX, pp. 51–56.
Mundhada, H., Miguel, J.S., Schneider, K., Koza, A., Christensen, H.B., Klein, T., Phaneuf, P.V., Herrgard, M., Feist, A.M., Nielsen, A.T., 2016. Increased production of L-serine in Escherichia coli through adaptive laboratory evolution. Metab. Eng. https://doi.org/10.1016/j.ymben.2016.11.008.
Mutalik, V.K., Guimaraes, J.C., Cambray, G., Mai, Q.-A., Christoffersen, M.J., Martin, L., Yu, A., Lam, C., Rodriguez, C., Bennett, G., Keasling, J.D., Endy, D., Arkin, A.P., 2013. Quantitative estimation of activity and quality for collections of functional genetic elements. Nat. Methods 10, 347–353.
Nyerges, Á., Csörgő, B., Draskovits, G., Kintses, B., Szili, P., Ferenc, G., Révész, T., Ari, E., Nagy, I., Bálint, B., Vásárhelyi, B.M., Bihari, P., Számel, M., Balogh, D., Papp, H., Kalapis, D., Papp, B., Pál, C., 2018. Directed evolution of multiple genomic loci allows the prediction of antibiotic resistance. Proc. Natl. Acad. Sci. USA. https://doi.org/10.1073/pnas.1801646115.
Oliphant, T., 2015. Guide to NumPy, 2nd ed. CreateSpace.
Orr-Weaver, T.L., Szostak, J.W., Rothstein, R.J., 1981. Yeast transformation: a model system for the study of recombination. Proc. Natl. Acad. Sci. USA 78, 6354–6358.
Oulmouden, A., Karst, F., 1990. Isolation of the ERG12 gene of Saccharomyces cerevisiae encoding mevalonate kinase. Gene 88, 253–257.
Özaydın, B., Burd, H., Lee, T.S., Keasling, J.D., 2013. Carotenoid-based phenotypic screen of the yeast deletion collection reveals new genes with roles in isoprenoid production. Metab. Eng. 15, 174–183.
Phillips, P.C., 2008. Epistasis–the essential role of gene interactions in the structure and evolution of genetic systems. Nat. Rev. Genet. 9, 855–867.
Plesa, C., Sidore, A.M., Lubock, N.B., Zhang, D., Kosuri, S., 2018. Multiplexed gene synthesis in emulsions for exploring protein functional landscapes. Science 359, 343–347.
Raman, S., Rogers, J.K., Taylor, N.D., Church, G.M., 2014. Evolution-guided optimization of biosynthetic pathways. Proc. Natl. Acad. Sci. USA 111, 201409523.
Raymond, S., Nicot, F., Jeanne, N., Delfour, O., Carcenac, R., Lefebvre, C., Cazabat, M., Sauné, K., Delobel, P., Izopet, J., 2017. Performance comparison of next-generation sequencing platforms for determining HIV-1 coreceptor use. Sci. Rep. 7, 42215.
Resnick, M.A., Martin, P., 1976. The repair of double-strand breaks in the nuclear DNA of Saccharomyces cerevisiae and its genetic control. Mol. Gen. Genet. 143, 119–129.
Ronda, C., Maury, J., Jakočiunas, T., Jacobsen, S.A.B., Germann, S.M., Harrison, S.J., Borodina, I., Keasling, J.D., Jensen, M.K., Nielsen, A.T., 2015. CrEdit: CRISPR mediated multi-loci gene integration in Saccharomyces cerevisiae. Microb. Cell Fact. 14, 97.
Ronda, C., Pedersen, L.E., Hansen, H.G., Kallehauge, T.B., Betenbaugh, M.J., Nielsen, A.T., Kildegaard, H.F., 2014. Accelerating genome editing in CHO cells using CRISPR Cas9 and CRISPy, a web-based target finding tool. Biotechnol. Bioeng. https://doi.org/10.1002/bit.25233.
Rouet, P., Smih, F., Jasin, M., 1994. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proc. Natl. Acad. Sci. USA 91, 6064–6068.
Roy, K.R., Smith, J.D., Vonesch, S.C., Lin, G., Tu, C.S., Lederer, A.R., Chu, A., Suresh, S., Nguyen, M., Horecka, J., Tripathi, A., Burnett, W.T., Morgan, M.A., Schulz, J., Orsley, K.M., Wei, W., Aiyar, R.S., Davis, R.W., Bankaitis, V.A., Haber, J.E., Salit, M.L., St Onge, R.P., Steinmetz, L.M., 2018. Multiplexed precision genome editing with trackable genomic barcodes in yeast. Nat. Biotechnol. 36, 512–520.
Ruff, P., Koh, K.D., Keskin, H., Pai, R.B., Storici, F., 2014. Aptamer-guided gene targeting in yeast and human cells. Nucleic Acids Res. 42, e61.
Ryan, O.W., Skerker, J.M., Maurer, M.J., Li, X., Tsai, J.C., Poddar, S., Lee, M.E., DeLoache, W., Dueber, J.E., Arkin, A.P., Cate, J.H.D., 2014. Selection of chromosomal DNA libraries using a multiplex CRISPR system. Elife e03703.
Schirmer, M., Ijaz, U.Z., D'Amore, R., Hall, N., Sloan, W.T., Quince, C., 2015. Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. Nucleic Acids Res. 43, e37.
Sikorski, R.S., Hieter, P., 1989. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. Genetics 122.
Skjoedt, M.L., Snoek, T., Kildegaard, K.R., Arsovska, D., Eichenberger, M., Goedecke, T.J., Rajkumar, A.S., Zhang, J., Kristensen, M., Lehka, B.J., Siedler, S., Borodina, I., Jensen, M.K., Keasling, J.D., 2016. Engineering prokaryotic transcriptional activators as metabolite biosensors in yeast. Nat. Chem. Biol. 12, 951–958.
Storici, F., Lewis, L.K., Resnick, M.A., 2001. In vivo site-directed mutagenesis using oligonucleotides. Nat. Biotechnol. 19, 773–776.
Storici, F., Resnick, M.a., 2003. Delitto perfetto targeted mutagenesis in yeast with oligonucleotides, In: Genetic Engineering. pp. 189–207.
Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J., Stahl, F.W., 1983. The double-strand-break repair model for recombination. Cell 33, 25–35.
Tarshis, L.C., Proteau, P.J., Kellogg, B.A., Sacchettini, J.C., Poulter, C.D., 1996. Regulation of product chain length by isoprenyl diphosphate synthases. Proc. Natl. Acad. Sci. USA 93, 15018–15023.
Verwaal, R., Wang, J., Meijnen, J.-P., Visser, H., Sandmann, G., van den Berg, J.A., van Ooyen, A.J.J., 2007. High-level production of beta-carotene in Saccharomyces cerevisiae by successive transformation with carotenogenic genes from Xanthophyllomyces dendrorhous. Appl. Environ. Microbiol. 73, 4342–4350.
Wang, H.H., Isaacs, F.J., Carr, P.A., Sun, Z.Z., Xu, G., Forest, C.R., Church, G.M., 2009. Programming cells by multiplex genome engineering and accelerated evolution. Nature 460, 894–898.
Wong, T.S., Tee, K.L., Hauer, B., Schwaneberg, U., 2004. Sequence saturation mutagenesis (SeSaM): a novel method for directed evolution. Nucleic Acids Res. 32, e26.
Wu, S., Letchworth, G.J., 2004. High efficiency transformation by electroporation of Pichia pastoris pretreated with lithium acetate and dithiothreitol. Biotechniques 36, 152–154.
Yang, J., Ruff, A.J., Arlt, M., Schwaneberg, U., 2017. Casting epPCR (cepPCR): a simple random mutagenesis method to generate high quality mutant libraries. Biotechnol. Bioeng. 114, 1921–1927.
Ye, L., Lv, X., Yu, H., 2016. Engineering microbes for isoprene production. Metab. Eng. 38, 125–138.